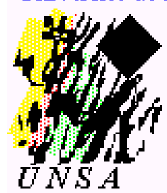
Université de Nice - Sophia Antipolis

Centre National de la Recherche Scientifique



U.M.R. 5651 "Espace" du C.N.R.S.

Equipe "Gestion et Valorisation de l'Environnement"



NOTICE D'UTILISATION DU LOGICIEL

HYDROLAB

(Version 98.2 pour EXCEL4, EXCEL5 et EXCEL97)

Ce logiciel a été développé par

J.P. LABORDE
Dr. es Sciences en Hydrologie
Ingénieur hydrogéologue de l'E.N.S.G.
Ingénieur hydraulicien de l'E.N.S.E.E.I.H.T.
Professeur à l'Université de Nice - Sophia Antipolis

et avec l'aide de

N. MOUHOUS
Ingénieur d'Etat en Hydraulique
Chercheuse à l'U.M.R. 5651

Octobre 1998

Equipe Gestion et Valorisation de l'Environnement de l'UM.R. 5651 du C.N.R.S.
98, Boulevard Edouard Herriot - B.P. 209 - 06204 NICE Cedex 3
Tél. 04 93 37 54 61 et 04 93 37 53 80 , Fax. 04 93 37 54 30, E. Mail : laborde@unice.fr

SOMMAIRE

UTILISATION PRATIQUE

1 Introduction

L'objet d'HYDROLAB n'est certainement pas de remplacer des logiciels beaucoup plus complets tels qu'on peut en trouver dans le commerce.

Notre volonté a été tout d'abord de créer un **outil très simple** parfaitement **intégré au logiciel EXCEL** dont l'utilisation est universelle. Pour se servir correctement d'HYDROLAB, il est indispensable d'avoir des bases sur l'utilisation d'EXCEL.

Il est préférable d'utiliser HYDROLAB avec un environnement utilisant la virgule comme séparateur décimal. Cependant, on peut également utiliser HYDROLAB avec le point comme séparateur ; dans ce cas de figure il faut par contre remplacer manuellement les virgules par des points dans les sorties des programmes de régression multiple, d'analyse en composantes principales et de comblement des lacunes. Il suffit de sélectionner l'ensemble de la feuille puis d'utiliser la fonction "Remplacer" du menu "Edition".

D'une façon générale, HYDROLAB est donc interfacé via EXCEL à la plus part des autres logiciels tels que : WORD, SURFER, …

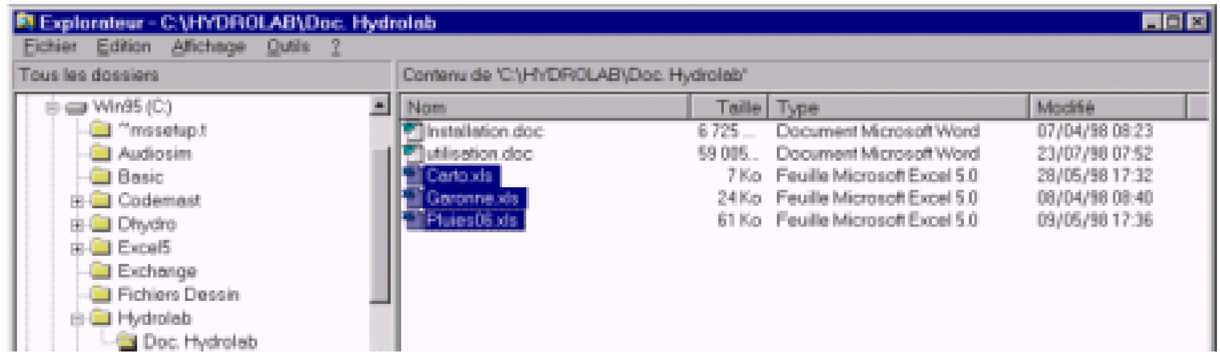
Le second objectif est de répondre aux **questions les plus fréquemment posées aux Hydrologues**. Ces questions portent essentiellement sur :

- L'analyse univariée (**Ajustements**) ; - L'analyse multivariée (**Régressions multiples**) ; - L'analyse en composantes principales (**ACP**) ; - Le **comblement de lacunes** dans des séries de données ; - La détection d'**anomalies** dans les séries de données ; - L'analyse spatiale (**variographie**).	Ces points sont traités par des **macros** commandes.
- Des **fonctions statistiques classiques** telles que F de Fisher-Snedecor, intégrale de Gauss… ; - Des fonctions liées à l'estimation de **l'évapo-transpiration potentielle** (Penmann, Durée du jour, radiation, Mc culloch, FAO,…) ; - Une fonction pour passer de **l'évapotranspiration** potentielle à la **réelle** ; - Des fonctions de passage des **coordonnées géographiques** à différentes coordonnées Lambert.	Ces points sont traités par des **"fonctions personnalisées"** au sens d'EXCEL.

Vous avez ici la version 98.2. Elle est l'aboutissement actuel d'une lignée de commandes écrites depuis 1990 dans les versions 2.2 à 97 d'EXCEL. Au fur et à mesure, HYDROLAB s'enrichit, se perfectionne et se corrige en fonction des utilisations qui en sont faites. Aussi n'hésitez pas à nous faire part des anomalies que vous auriez détectées, des améliorations que vous proposeriez et des développements que vous souhaiteriez, à l'adresse suivante :

UPRESA 6046 du C.N.R.S.
98 Bd E. Herriot -B.P. 209
06204 NICE Cedex 3
Fax; 04 93 37 54 30
E. mail : laborde@unice.fr

Les données utilisées à titre d'exemple dans ce guide sont fournies dans classeur sous la forme des trois fichiers EXCEL suivants :



2 Les macros commandes

2.1 Ajustements

HYDROLAB permet de réaliser des ajustements aux lois les plus fréquemment utilisées en hydrologie :

- Loi de Gauss (loi normale) ;
- Loi de Galton (loi log-normale) ;
- Loi Racine-normale ;
- Loi de Gumbel ;
- Loi de Fuller ;
- Loi de Weibull ;
- Loi de Poisson.

2.1.1 Règles communes :

2.1.1.1 Vos données sont préservées

D'une façon systématique, les feuilles Excel où sont situées vos données de départ ne sont jamais modifiées. Les feuilles de commandes et de présentation des résultats utilisés par HYDROLAB sont verrouillées. Vous pouvez les modifier mais vous ne pouvez pas enregistrer les modifications par erreur. Au cas ou vous auriez modifier une feuille originale d'HYDROLAB, il suffit de la fermer puis de la rouvrir.

Si vous voulez conserver vos résultats pour les faire figurer dans un rapport, il suffit de copier ce qui vous intéresse dans les feuilles Excel puis de les copier sous Word.

Si vous voulez enregistrer les résultats sous forme de fichier Excel, il convient d'utiliser la procédure "Enregistrer sous" du menu "Fichier". Vous pouvez alors choisir librement les noms des feuilles Excel à enregistrer. Attention ! Il est préférable d'enregistrer en premier les feuilles de graphique (extension .xlc) puis seulement les feuilles de calcul (extension .xls).

2.1.1.2 Sélection de la plage de données

L'échantillon que l'on veut ajuster doit être nécessairement une suite de valeurs numériques (ou des formules numériques) situées sur une seule et unique colonne d'une feuille EXCEL. Si ce n'était pas le cas vous devez vous y ramener par les différentes fonctionnalités d'EXCEL.

Il suffit tout d'abord de sélectionner la plage de cellules à traiter. Sur la figure 8 on s'apprête à traiter les 17 pluies annuelles à Cannes de 1959 à 1976. Sur la figure 9 nous avons une autre sélection tout à fait valide concernant des pluies à Antibes de 1960 à 1974 à l'exclusion de l'année 1968 pour laquelle il y' a une lacune. Pour obtenir cette sélection multiple, il a suffit de sélectionner la plage 1127 à 358 mm puis de maintenir enfoncée la touche Ctrl et sélectionner la plage 768 à 1075 mm. Par contre la sélection illustrée par la figure 10 n'est pas valide puisqu'on a pris en compte un terme alphanumérique correspondant à la lacune d'observation de 1962 à Clans.

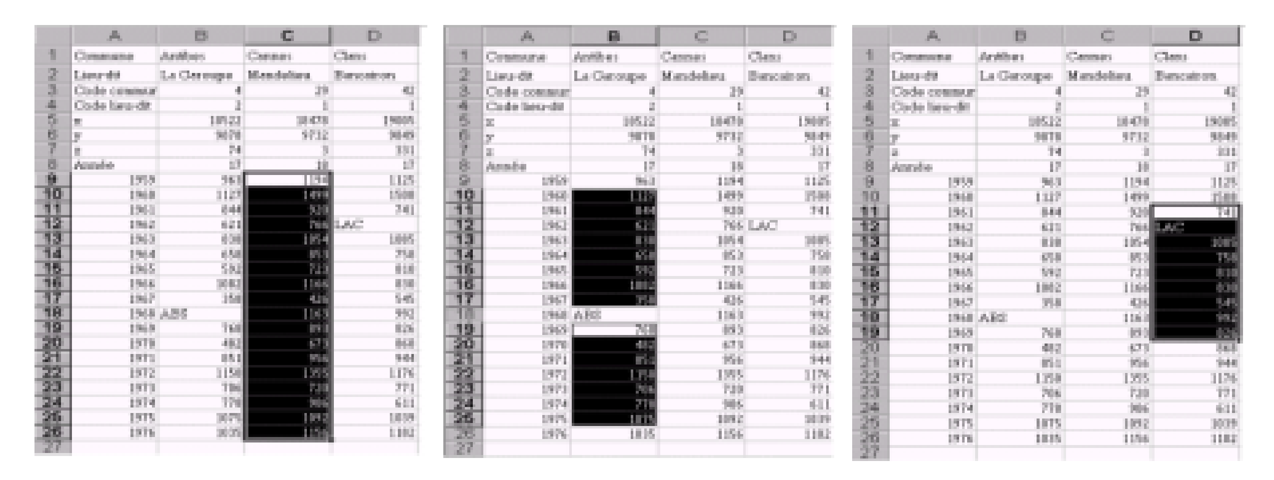


Fig.8 : Sélection continue valide *Fig.9 : Sélection discontinue valide* *Fig. 10 : Sélection non valide*

2.1.1.3 Lancement de l'ajustement

Une fois les données sélectionnées, il suffit pour procéder à un ajustement, de choisir dans la barre de menu HYDROLAB, le type de loi désiré.

Dans l'exemple de la figure 11 les pluies à Cannes de 1959 à 1976 seront ajustées à une loi Normale.

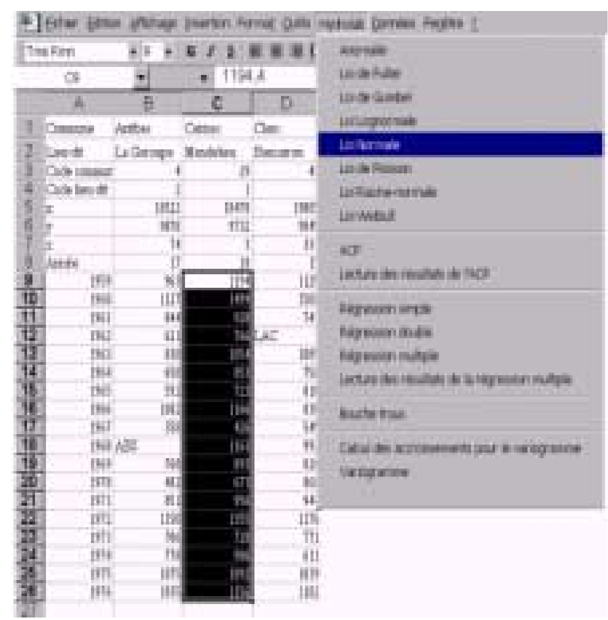


Fig.11 : Lancement d'un ajustement

2.1.1.4 Les feuilles de résultats

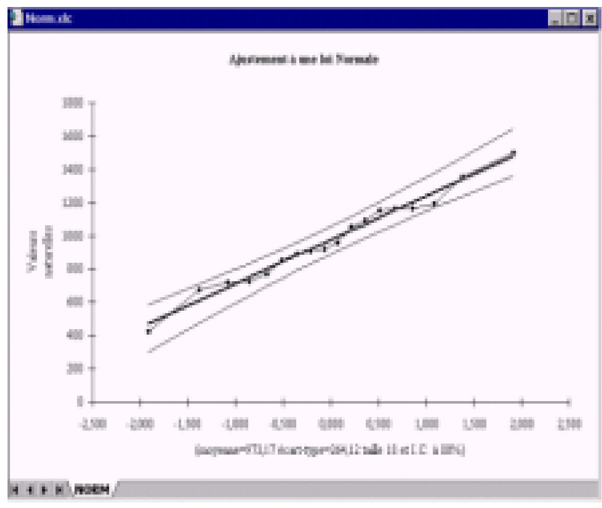


Fig. 12 : Exemple de résultats graphiques

Quelques instants après le lancement, vous verrez apparaître les feuilles de résultats. Elles sont toujours de même type et superposées. Celle qui vous est visible, contient la représentation graphique de l'ajustement et les paramètres essentiels. Comme l'indique la figure 12, les feuilles de graphiques portent l'extension ".xlc" (ici précédée de Norm puisqu'il s'agit d'une loi de Gauss).

Les valeurs de la variable sont toujours portées en ordonnées et les fréquences au non dépassement en abscisses, selon les échelles adaptées à la loi.
La loi théorique est matérialisée par une droite en trait fort. La plus part du temps vous verrez également deux courbes en trait fin limitant un intervalle de confiance à 80% (intervalle que vous pourrez modifier).

On trouve également des informations succinctes sur la taille de l'échantillon traité, les paramètres ajustés et l'intervalle de confiance tracé.

On peut toujours habiller ce graphique en utilisant les fonctionnalités d'EXCEL En cliquant sur le titre ou les axes on peut en modifier les textes, les polices … Par "affichage", puis "barre d'outil" on peut accéder à toute forme de traits, de motifs … La figure 13 en est un exemple. Il suffit lorsque le graphique vous satisfait, de le copier (Ctrl C) puis de la coller par exemple dans votre rapport sous WORD.

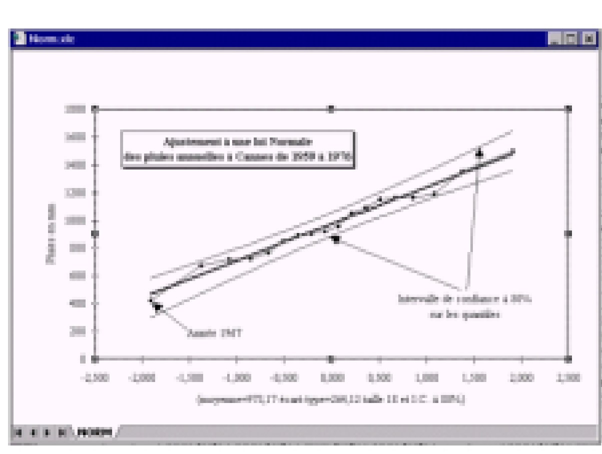
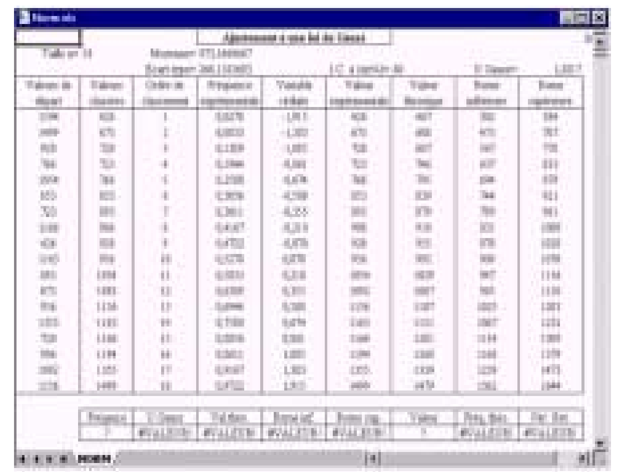


Fig. 13 : Exemple d'habillage personnalisé

Une deuxième feuille est située sous le graphique, vous pouvez y accéder soit par le menu "fenêtre" où apparaît le nom d'une feuille de calcul avec le suffixe ".xls", soit en diminuant la taille de la fenêtre du graphique. On peut bien sur avoir les deux fenêtres visibles en mosaïque.



Ajustement à une loi de Gauss

Taille n= 18 Moyenne= 973,1666667

I.C. à (en%)= 80

Fréquence	U Gauss	Val.théo.	Borne inf.	Borne sup.	Valeur	Fréq. théo.	Per. Ret.
?	#VALEUR!	#VALEUR!	#VALEUR!	#VALEUR!	?	#VALEUR!	#VALEUR!

Fig. 14 : Exemple de résultats numériques

Les premières colonnes sont toujours les mêmes, les valeurs dans l'ordre de départ, les valeurs rangées par ordre croissant, l'ordre de classement, la fréquence expérimentale calculée par la formule :

$$\hat{F}(x_i) = \frac{i - 0.5}{n}$$ (i rang de classement, n taille de l'échantillon)

En y trouve également, les paramètres de l'ajustement, la taille de l'échantillon et l'intervalle de confiance exprimé en pour-cent. On peut modifier la valeur de 80% par défaut en remplaçant simplement la valeur 80 par celle souhaitée.

Dans le bas de la feuille, on trouve un second tableau qui permet d'obtenir des valeurs quantiliques ou au contraire trouver la fréquence et la période de retour d'une valeur quelconque. Il suffit de mettre la valeur souhaitée à la place des points d'interrogation. Dans le cas où l'on désirerait un tableau comportant plusieurs estimations de ce type, il suffit de sélectionner les cellules comprises entre celle sous "Fréquence" et celle sous "Per. Ret." , puis de les recopier vers le bas sur autant de lignes que nécessaire.

Ajustement à une loi de Gauss

Taille n= 18 Moyenne= 973,1666667

Ecart-type= 266,1153683 I.C. à (en%)= 70 U Gauss= 1,0364

Valeurs de départ	Valeurs classées	Ordre de classement	Fréquence expérimentale	Variable réduite	Valeur expérimentale	Valeur théorique	Borne inférieure	Borne supérieure
1194	426	1	0,0278	-1,915	426	467	380	568
1499	673	2	0,0833	-1,383	673	608	504	699
920	728	3	0,1389	-1,085	728	687	584	762
766	723	4	0,1944	-0,861	723	746	661	817
1058	766	5	0,2500	-0,674	766	795	716	863
853	853	6	0,3056	-0,508	853	839	764	908
723	893	7	0,3611	-0,355	893	879	808	945
1166	906	8	0,4167	-0,210	906	918	849	983
426	928	9	0,4722	-0,070	928	955	888	1021
1163	956	10	0,5278	0,070	956	992	926	1059
893	1054	11	0,5833	0,210	1054	1029	963	1098
673	1092	12	0,6389	0,355	1092	1067	1001	1138
956	1156	13	0,6944	0,508	1156	1107	1041	1182
1355	1163	14	0,7500	0,674	1163	1151	1083	1230
728	1166	15	0,8056	0,861	1166	1201	1130	1286
906	1194	16	0,8611	1,085	1194	1260	1185	1352
1092	1355	17	0,9167	1,383	1355	1339	1257	1443
1156	1499	18	0,9722	1,915	1499	1479	1383	1606

Fréquence	U Gauss	Val.théo.	Borne inf.	Borne sup.	Valeur	Fréq. théo.	Per. Ret.
0,8	0,841	1195	1125	1280	400	0,013	1,0
0,9	1,282	1312	1232	1412	800	0,256	1,3
0,95	1,645	1408	1319	1523	1200	0,805	5,1

Dans le tableau ci-contre, l'intervalle de confiance a été ramené à 70% et l'on a évalué les valeurs de période de retour 5, 10 et 20 ans, ainsi que les périodes de retour de pluies de 400, 800 et 1200 mm.

Fig. 15 : Modification d'une feuille de résultats numériques

2.1.2 La loi de Gauss

La mise en œuvre de la loi de Gauss correspond intégralement au dispositif commun de lancement et de présentation des résultats. Cette loi est définie sur $[-\infty, +\infty]$ et peut convenir à n'importe quel échantillon.

Il nous parait seulement bon de préciser quelques points techniques. La fonction de répartition $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}}$ du avec $u = \frac{x - \overline{x}}{\sigma}$ dépends de deux paramètres qui sont la moyenne $\overline{x}$ et l'écart-type σ.

• Ces paramètres moyenne et écart type sont évalués selon les estimateurs sans biais suivant :

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \sigma = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1}}$$

(paramètres estimés par la méthode des moments)

• L'intégrale de Gauss est approchée par les formules suivantes :

Calcul de u pour F <= 0,5 :

- u = x **- Erreur !**avec x = **Erreur !**

(pour F > 0,5 : on fait F = 1 **-** F et u = **-** u)

Calcul de F pour u > 0 :

$$F(u) = 1 - \frac{0,5}{(1 + 0,196854\, x^1 + 0,115194\, x^2 + 0,000344\, x^3 + 0,019527\, x^4)^4}$$

(pour u < 0 on fera u = **-** u et F = 1 **-** F)

• pour l'intervalle de confiance à α % sur le quantile x_F les bornes sont évaluées par :

x_F ± **Erreur !**σ *(Informations Techniques du CTGREF, Cahier 31, N°2, 1978)*

(2 fois plus pour la borne supérieure et 2 fois moins pour la borne inférieure)

t_F : variable réduite de Gauss ayant la fréquence au non-dépassement F

t_α : variable réduite de Gauss ayant la fréquence au non-dépassement $1 - \frac{1-\alpha}{2}$.

2.1.3 La loi de Galton (loi Log-normale)

La mise en œuvre de la loi de Galton correspond intégralement au dispositif commun de lancement et de présentation des résultats. Il faut cependant préciser quelques points techniques :

La fonction de répartition est $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}} du$ avec $u = \frac{\log(x - x_o) - a}{b}$; ici nous considérerons que **c'est la variable log(x-x_o) qui suit une loi de Gauss** à deux paramètres :

- a : la moyenne des log(x-x_o) ;
- b : l'écart-type des log(x-x_o) .

$$a = \frac{\sum_{i=1}^{n} \log(x_i - x_0)}{n} \qquad b = \sqrt{\frac{\sum_{i=1}^{n} \log^2(x_i - x_0) - n\,a^2}{n-1}}$$

Dans cette approche on voit bien que le paramètre de position x_o n'est pas considéré comme un paramètre d'ajustement mais comme une caractéristique connue a priori (borne inférieure de l'intervalle de définition). Par défaut, x_o est supposé nul, mais vous pouvez lui donner une autre valeur en l'entrant directement dans la cellule située à droite de " x_o =". Il faut évidemment que x_o soit strictement inférieur à la plus petite valeur de l'échantillon.

Ajustement à une loi Log-normale

Taille n= 12 — Moy log(x-xo)= 1,511708719

xo= 0 — E.T. log(x-xo)= 0,354917709 — I.C. à (en%)= 80

Valeurs de départ	Valeurs classées	Ordre de classement	Fréquence expérimentale	Variable réduite	Valeur expérimentale	Valeur théorique
12	12	1	0,042	-1,732	12	7,887977
15	15	2	0,125	-1,150	15	12,687929
32	16	3	0,208	-0,812	16	16,730857
56	21	4	0,292	-0,548	21	20,756851
21	24	5	0,375	-0,318	24	25,04759
45	29	6	0,458	-0,104	29	29,829879
265	32	7	0,542	0,104	32	35,379358
48	32	8	0,625	0,318	32	42,134271
16	45	9	0,708	0,548	45	50,844521
24	48	10	0,792	0,812	48	63,08179
29	56	11	0,875	1,150	56	83,178426
32	265	12	0,958	1,732	265	133,79374

Ajustement à une loi Log-normale

Taille n= 12 — Moy log(x-xo)= 1,414197921

xo= 5 — E.T. log(x-xo)= 0,412269895 — I.C. à (en%)= 80

Valeurs de départ	Valeurs classées	Ordre de classement	Fréquence expérimentale	Variable réduite	Valeur expérimentale	Valeur théorique
12	12	1	0,042	-1,732	12	10,813327
15	15	2	0,125	-1,150	15	13,78779
32	16	3	0,208	-0,812	16	17,806682
56	21	4	0,292	-0,548	21	20,424733
21	24	5	0,375	-0,318	24	24,187293
45	29	6	0,458	-0,104	29	28,505884
265	32	7	0,542	0,104	32	33,657219
48	32	8	0,625	0,318	32	40,106864
16	45	9	0,708	0,548	45	48,669498
24	48	10	0,792	0,812	48	61,101286
29	56	11	0,875	1,150	56	82,354812
32	265	12	0,958	1,732	265	138,35995

Fig. 16 : Modification du paramètre de position dans une loi de Galton
(à gauche x_o est pris à priori égal à 0, à droite on s'est imposé x_o = 5 en le tapant directement dans la feuille)

Pour le calcul des quantiles et de leurs intervalles de confiance, les calculs sont faits sur la variable transformée $\log(x-x_0)$ supposée gaussienne, puis ramenés en variable naturelle par $x = x_0 + 10^{\log(x-x0)}$.

2.1.4 La loi Racine-normale

La mise en œuvre de la loi de Racine-normale correspond intégralement au dispositif commun de lancement et de présentation des résultats. Comme pour la loi de Galton on se ramène à une loi de Gauss par un changement de variable :

La fonction de répartition est $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{u^2}{2}} du$ avec $u = \frac{\sqrt{x} - a}{b}$; ici nous considérerons que **c'est la variable $\sqrt{x}$ qui suit une loi de Gauss** à deux paramètres :

- a : la moyenne des $\sqrt{x}$;
- b : l'écart-type des $\sqrt{x}$.

$$a = \frac{\sum_{i=1}^{n} \sqrt{x_i}}{n} \qquad b = \sqrt{\frac{\sum_{i=1}^{n} x_i - n a^2}{n-1}}$$

Cette loi est définie sur $[0,+\infty]$ et ne peut donc convenir qu'à une population de valeurs positives. Pour le calcul des quantiles et de leurs intervalles de confiance, les calculs sont faits sur la variable transformée $\sqrt{x}$ supposée gaussienne, puis ramenés en variable naturelle par $x = \sqrt{x}^{\,2}$.

2.1.5 La loi de Gumbel

La mise en œuvre de la loi de Gumbel correspond intégralement au dispositif commun de lancement et de présentation des résultats. La fonction de répartition est :

$$F(x) = e^{e^{-u}} \quad \text{avec} \quad u = \frac{x - x_0}{g} \text{ (variable réduite de Gumbel)}$$

Le terme u est la variable réduite de Gumbel ; x_0 est le paramètre de position (mode) et g est le paramètre d'échelle différent de zéro et positif (g est aussi appelé "gradex").

Cette loi est définie sur $[-\infty,+\infty]$ et peut convenir à n'importe quel échantillon.

On peut noter dès à présent le **comportement asymptotiquement exponentiel** de la distribution de Gumbel : si F tend vers 1, en posant T = **Erreur !**, la variable réduite de Gumbel u tend vers Ln T.

Les paramètres sont estimés par la méthode des moments :

$$g = 0{,}78\ \sigma \text{ et } \quad x_0 = \overline{x,} \ - 0{,}577\ s$$

L'intervalle de confiance à α % sur un quantile x_F s'exprime en fonction de l'écart-type σ par :

$$\hat{x}_F - h_1\,\sigma \le x_F < \hat{x}_F + h_2\,\sigma$$

où h_1 et h_2 sont des paramètres dépendant de la taille n de l'échantillon de la fréquence F et de la valeur de α.

• h_1 et h_2 seront évalués par la formule suivante (avec le signe + pour h_2 et le signe - pour h_1) :

$$h_{1,2} = \frac{\frac{u_\alpha}{n}\sqrt{1 + 1{,}13 t_F + 1{,}1 t_F^2} \pm \frac{u_\alpha^2}{n}(1{,}1 t_F + 0{,}57)}{1 - 1{,}1\frac{u_\alpha^2}{n}}$$

(Informations Techniques du CTGREF, Cahier 31, N°2, 1978)

• u_α est la variable réduite de Gauss correspondant à la fréquence au non-dépassement $1-\frac{1-\alpha}{2}$

• t_F est la variable réduite de Gumbel correspondant à la fréquence au non-dépassement F, ramenée à sa moyenne et à son écart-type :

t_F = **Erreur !**

2.1.6 La loi de Fuller

La mise en œuvre de la loi de Fuller correspond intégralement au dispositif commun de lancement et de présentation des résultats. La fonction de répartition est :

$$F(x) = 1 - e^{-\frac{x-x_0}{g}}$$

Cette loi a donc deux paramètres (x_0 et g) et est définie sur l'intervalle $[x_0, +\infty]$. Nous considérons ici que **x_0 est la borne inférieure** de l'intervalle de définition et qu'il est donc **connu a priori**. Dans la feuille la valeur par défaut de x_0 est la plus petite valeur de l'échantillon mais on peut toujours se donner une autre valeur de x_0 en la tapant directement dans la cellule située à droite de " x_0=". Il faut bien sûr, que x_0 soit inférieur ou égal à la plus petite valeur de l'échantillon.

Ajustement à une loi de Fuller des valeurs supérieures à un seuil

Taille n= 10 Moyenne= 34,3 Borne inf. xo= 4 U Gauss = 1,2817288

Ecart-type= 44,69911757 Gradex g= 30,3 I.C. à (en %) = 80

Valeurs de départ	Valeurs classées	Ordre de classement	Fréquence expérimentale	Pseudo période de retour	Valeurs expérimentales	Valeurs théoriques	Borne inférieure	Borne supérieure
4	4	1	0,050	1,1	4	1,75956	1,203002	2,3157185
5	5	2	0,150	1,2	5	5,5743993	3,811619	7,3371791
12	12	3	0,250	1,3	12	9,8678951	6,747119	12,987871
25	14	4	0,350	1,5	14	14,775854	18,10332	19,448389
14	16	5	0,450	1,8	16	20,585889	14,8211	26,990315
16	25	6	0,550	2,2	25	27,388814	18,72771	36,049917
38	32	7	0,650	2,9	32	36,008899	24,62189	47,395913
156	38	8	0,750	4,0	38	47,549897	32,5131	62,586494
41	41	9	0,850	6,7	41	65,871215	44,49389	85,648541
32	156	10	0,950	20,0	156	102,75362	78,26007	135,24716

Ajustement à une loi de Fuller des valeurs supérieures à un seuil

Taille n= 10 Moyenne= 34,3 Borne inf. xo= 0 U Gauss = 1,2817288

Ecart-type= 44,69911757 Gradex g= 34,3 I.C. à (en %) = 80

Valeurs de départ	Valeurs classées	Ordre de classement	Fréquence expérimentale	Pseudo période de retour	Valeurs expérimentales	Valeurs théoriques	Borne inférieure	Borne supérieure
4	4	1	0,050	1,1	4	1,75956	1,203002	2,3157185
5	5	2	0,150	1,2	5	5,5743993	3,811619	7,3371791
12	12	3	0,250	1,3	12	9,8678951	6,747119	12,987871
25	14	4	0,350	1,5	14	14,775854	18,10332	19,448389
14	16	5	0,450	1,8	16	20,585889	14,8211	26,990315
16	25	6	0,550	2,2	25	27,388814	18,72771	36,049917
38	32	7	0,650	2,9	32	36,008899	24,62189	47,395913
156	38	8	0,750	4,0	38	47,549897	32,5131	62,586494
41	41	9	0,850	6,7	41	65,871215	44,49389	85,648541
32	156	10	0,950	20,0	156	102,75362	78,26007	135,24716

Fig. 17 : Modification du paramètre de position dans une loi de Fuller
à gauche x_o est pris à priori égal à 4(plus petite valeur de l'échantillon),
à droite on s'est imposé x_o =0 en le tapant directement dans la feuille

Le paramètre g (que l'on peut également appeler gradex) est estimé par : $g = \overline{x} - x_0$

L'intervalle de confiance à α % sur un quantile x_F s'exprime ainsi :

$$\hat{x}_F - u_\alpha \frac{(\hat{x}_F - x_0)}{\sqrt{n}} < x_F < \hat{x}_F + u_\alpha \frac{(\hat{x}_F - x_0)}{\sqrt{n}}$$

• u_α est la variable réduite de Gauss correspondant à la fréquence au non-dépassement $1-\frac{1-\alpha}{2}$

(formule simplifiée d'après MIQUEL 1984 et valable si g est le seul paramètre ajusté).

2.1.7 La loi de Weibull

La fonction de répartition de Weibull est:

$$F(x) = 1 - e^{-\rho (x - x_o)^p}$$

... dont la densité est :

$$f(Q) = \rho\, p\, (x - x_o)^{p-1}\, e^{-\rho (x - x_o)^p}$$

Dans cette loi figurent trois paramètres ρ, p et x_o, mais ce dernier, borne inférieure de l'intervalle de définition, est choisi *a priori*. Il n'y a donc que deux paramètres à ajuster : ρ et p.

Les paramètres sont estimés par la méthode du maximum de vraisemblance :

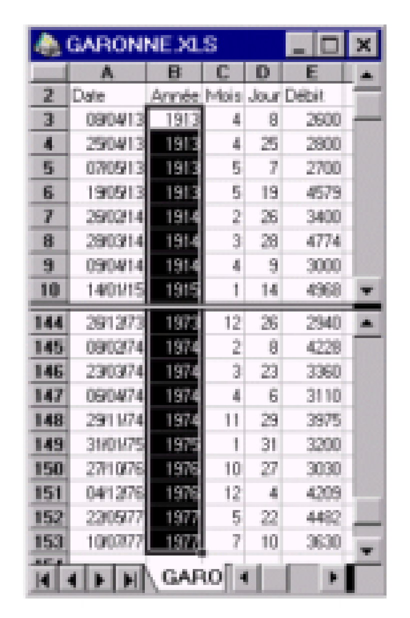
$$\frac{1}{p} = \frac{\sum_{j=1}^{nc} (x_j - x_o)^p \, \mathrm{Ln}(x_j - x_o)}{\sum_{j=1}^{nc} (x_j - x_o)^p} - \frac{1}{nc}\sum_{j=1}^{nc} \mathrm{Ln}(x_j - x_o) \quad \text{et} \quad \rho = \frac{nc}{\sum_{j=1}^{nc} (x_j - x_0)^p}$$

La première équation est résolue par itérations successives, puis on résout la seconde. Pour cette distribution, les intervalles de confiances ne sont pas donnés.

2.1.8 La loi de Poisson

La mise en œuvre de l'ajustement à une loi de Poisson se distingue nettement de la procédure commune aux autres lois.

Cette loi est utilisée principalement en hydrologie pour modéliser le nombre par an d'événements dépassant un seuil. Généralement les données se présentent donc sur la forme de deux colonnes : une colonne donnant les dates d'occurrence, la seconde les valeurs prises par la variable. Sur la figure 18, on trouve les données concernant les crues de la Garonne de 1913 à 1977, ayant un débit supérieur à 2500 m^3/s. On se propose d'ajuster à une loi de Poisson le nombre de crue par an ayant dépassé ce seuil (4 en 1913, 3 en 1914, … , 4 en 1974, 1 en 1975, 2 en 1976 et 2 en 1977). Il suffit de sélectionner les années où se sont produites ces crues (fig.18) puis de lancer la commande "loi de Poisson" du menu "HYDROLAB".



GARONNE.XLS

	A	B	C	D	E
2	Date	Année	Mois	Jour	Débit
3	08/04/13	1913	4	8	2600
4	25/04/13	1913	4	25	2800
5	07/05/13	1913	5	7	2700
6	19/05/13	1913	5	19	4579
7	26/02/14	1914	2	26	3400
8	28/03/14	1914	3	28	4774
9	09/04/14	1914	4	9	3000
10	14/01/15	1915	1	14	4968
144	26/12/73	1973	12	26	2940
145	08/02/74	1974	2	8	4228
146	23/03/74	1974	3	23	3360
147	06/04/74	1974	4	6	3110
148	29/11/74	1974	11	29	3975
149	31/01/75	1975	1	31	3200
150	27/10/76	1976	10	27	3030
151	04/12/76	1976	12	4	4209
152	22/05/77	1977	5	22	4482
153	10/07/77	1977	7	10	3630

GARO

Fig. 18 : sélection des occurrences

POIS.XLS

Ajustement à une loi de Poisson

Nombre de modalités rencontrées : 57

Nombre de modalités possibles : ?

Moyenne : #VALEUR! Ecart-type théo : #VALEUR! Ecart-type exp : #VALEUR!

Modalités Rencontrées	Modalité	Effectif	Effectif	Nb. exp.	Nb. théo.
1913	1913	4	0	#VALEUR!	#VALEUR!
1913	1914	3	1	16	#VALEUR!
1913	1915	5	2	15	#VALEUR!
1913	1916	2	3	8	#VALEUR!
1914	1917	2	4	11	#VALEUR!
1914	1918	3	5	5	#VALEUR!
1914	1919	3	6	2	#VALEUR!
1915	1920	4			

POIS

Fig. 19 : introduction manuelle du nombre de modalité

On obtient alors la feuille "POIS.XLS" dans laquelle vous devez introduire manuellement à la place du signe ?, le nombre de modalités possibles (voir figure 19). En effet, il n'est pas évident que le nombre de

modalités possibles soit 65 (1977-1913+1). Il se peut que la série présente des lacunes (nombre inférieur à 65), ou au contraire on peut imaginer que la série va au-delà de 1977 mais qu'il n'y a pas eu de dépassement du seuil (nombre supérieur à 65). Dés que l'on a rentré ce nombre de modalités, les feuilles de résultats se mettent à jour.

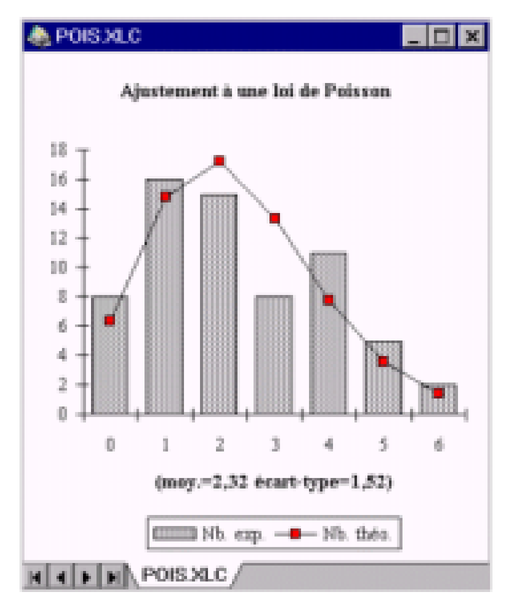
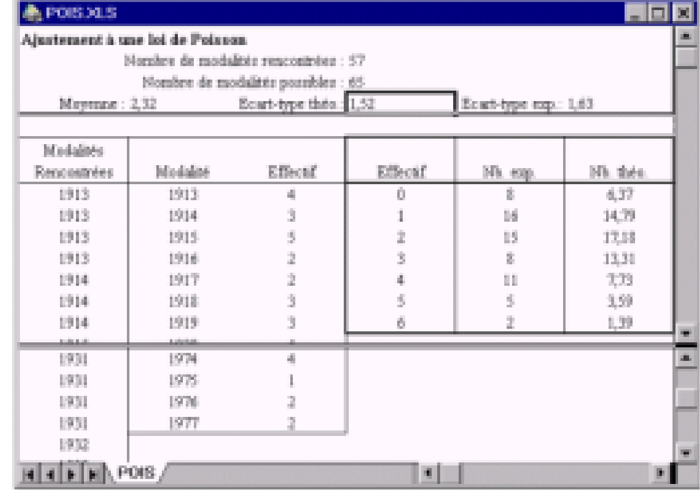


Fig. 20 : Feuille de résultats graphiques



Modalités Rencontrées	Modalité	Effectif	Effectif	Nb. exp.	Nb. théo.
1913	1913	4	0	8	6,37
1913	1914	3	1	16	14,79
1913	1915	5	2	15	17,18
1913	1916	2	3	8	13,31
1914	1917	2	4	11	7,73
1914	1918	3	5	5	3,59
1914	1919	3	6	2	1,39
1931	1974	4			
1931	1975	1			
1931	1976	2			
1931	1977	2			
1932					

Fig. 21 : Feuille de résultats numériques

Sur le plan technique la loi de Poisson (ou loi des événements rares) indique que la probabilité de rencontrer k événements est donnée par :

$$\text{Proba}(k) = \frac{e^{-\mu}\ \mu^{k}}{k!}$$

μ est le nombre moyen d'événement (ici 2,32=151/65) et on peut montrer que l'écart-type de k doit être égal à $\sqrt{\mu}$ (soit ici $1{,}52 = \sqrt{2{,}32}$, ce qui n'est effectivement pas éloigné de l'écart-type expérimental 1,63).

2.2 Régressions simples

2.2.1 Sélection des données

Les variables à traiter sont obligatoirement situées dans deux colonnes, la première correspondant à la variable à expliquer et la seconde à la variable explicative (si dans votre tableur cet ordre n'est pas respecté, il convient de le changer). Vous devez sélectionner uniquement les données numériques. La figure 22 vous donne trois exemples de sélection.

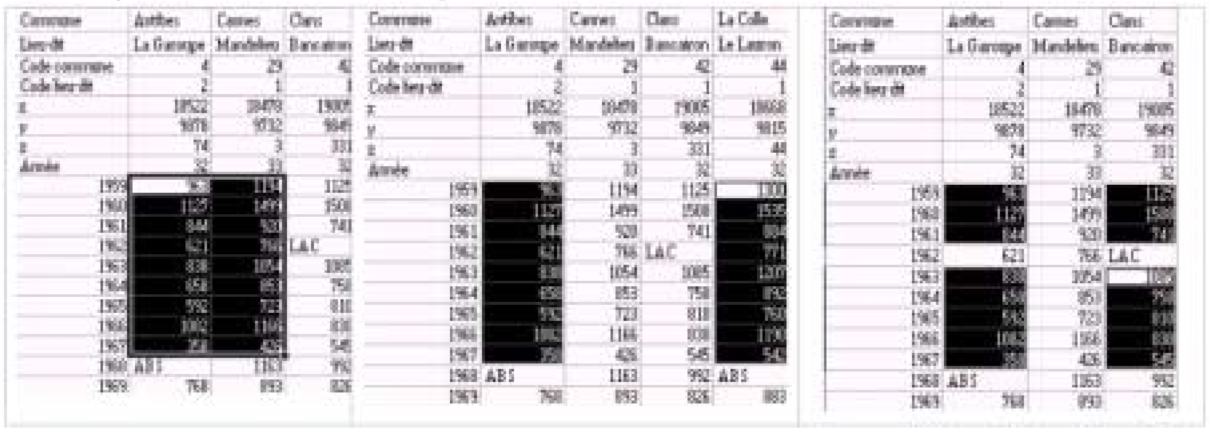


Fig. 22 : Exemples de sélection pour une régression simple

2.2.2 *Exécution de la macro*

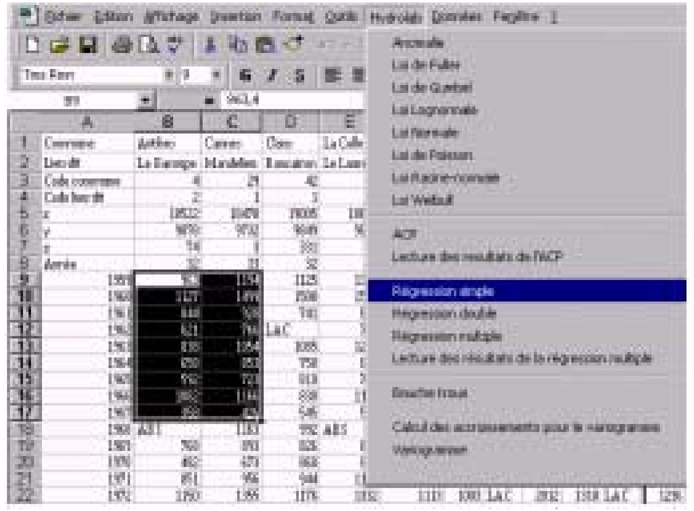


Fig. 23 : Lancement d'une régression simple

La figure 23 illustre le lancement d'une régression simple entre les pluies d'Antibes et Cannes sur la période commune 1959 - 1967.

Une fois la macro lancé, vous trouverez deux feuilles analogues à celles décrites pour les ajustements.

Une feuille intitulée "Regs.xlc" présente les points expérimentaux, la droite de régression et l'intervalle de confiance sur l'estimation de la variable à expliquer (première colonne et donc ici Antibes) en fonction de la variable explicative (figure 24).

Les valeurs numériques sont données dans la feuille "Regs.xls" comme l'illustre la figure 25.

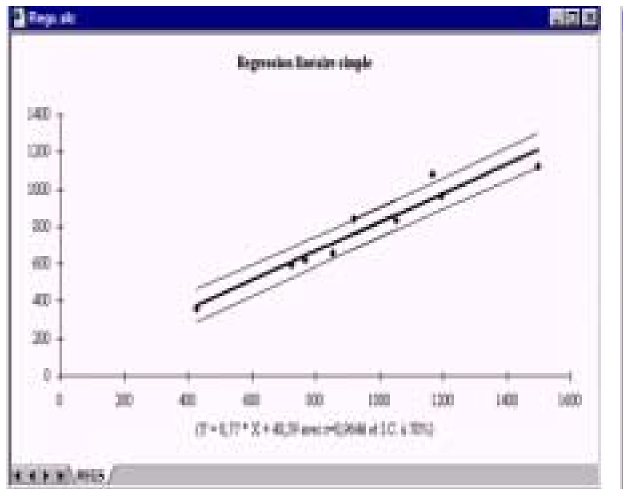


Fig. 24 : Graphique de la régression simple

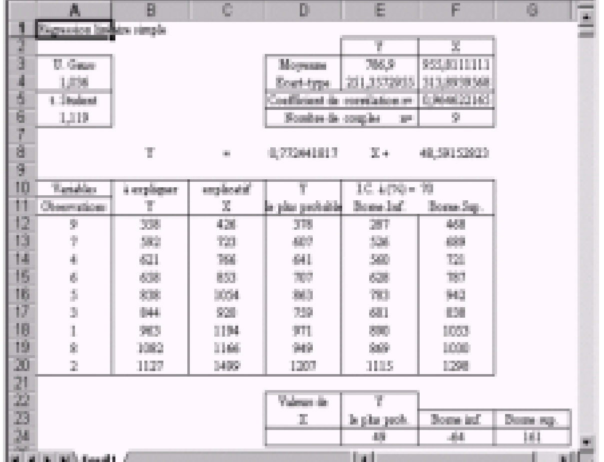


Fig. 25 : Valeurs numériques de la régression simple

Soit $\bar{x}_k$, $\bar{y}_k$, ${}_k\sigma_x$ et ${}_k\sigma_y$ les moyennes et écarts-types déterminés d'après les k observations communes ainsi que le coefficient de corrélation correspondant.

$$\bar{x}_k = \frac{\sum_{i=1}^{k} x_i}{k} \qquad \bar{x}_k = \frac{\sum_{i=1}^{k} y_i}{k} \qquad {}_k\sigma_x = \sqrt{\frac{\sum_{i=1}^{k} x_i^2 - k\bar{x}_k^2}{k-1}} \qquad {}_k\sigma_y = \sqrt{\frac{\sum_{i=1}^{k} y_i^2 - k\bar{y}_k^2}{k-1}} \qquad \rho = \frac{\sum_{i=1}^{k} x_i y_i - k\bar{x}_k\bar{y}_k}{(k-1)\,{}_k\sigma_x\,{}_k\sigma_y}$$

La régression de y en x s'écrit :

$$\bar{y}_{x_j} = \rho_k \frac{{}_k\sigma_y}{{}_k\sigma_x}\left(x_j - \bar{x}_k\right) + \bar{y}_k$$

$\bar{y}_{x_i}$ représentant évidemment la moyenne conditionnelle des y liés par x_j et si les hypothèses de normalité sont admissibles, $\bar{y}_{x_i}$ est la valeur la plus proche de y pour x_j donné.

Les bornes y_1 et y_2 de l'intervalle de confiance à $\alpha\%$ sur l'estimation de y à partir d'une réalisation x_i de la variable explicative sont données par la relation suivante :

$$y_{1,2} = \bar{y}_{x_i} \pm \sqrt{\left(1-\rho^2\right)\frac{(k-1)^2}{k(k-2)}}\sqrt{1+\frac{1}{k}+\frac{(x_i-\bar{x}_k)^2}{{}_k\sigma_x^2(k-1)}}$$

(formule tirée de YEVJEVICH 1982)

Un tableau situé en bas de la feuille "Regs.xls" permet d'évaluer les valeurs de y connaissant x.

2.2.3 Personnalisation des résultats

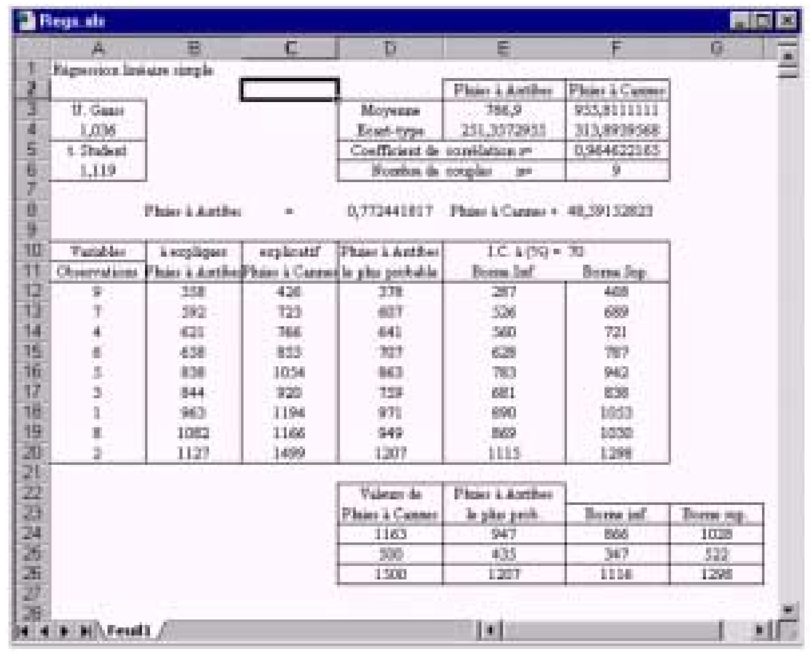


Fig. 26 : Personnalisation des résultats

Comme le montre la figure 26, il suffit de reporter (taper au clavier ou coller) dans les cellules B11 et C11 les noms des variables à expliquer et explicative pour que toute l'interface utilisateur en tienne compte. Dans le tableau du bas, on a cherché à évaluer la pluie à Antibes pour l'année 1968 où il était tombé 1163 mm à Cannes et ce qu'il en serait à Antibes pour des pluies à Cannes de 500 et 1500 mm. On aurait pu également changer la borne de l'intervalle de confiance en remplaçant le 70% de la cellule F10 pour toute autre valeur qu'on aurait souhaitée.

2.3 Régression double

2.3.1 Sélection des données

Les données sont disposées comme pour la régression simple mais sur trois colonnes ; la première colonne contenant toujours la variable à expliquer.

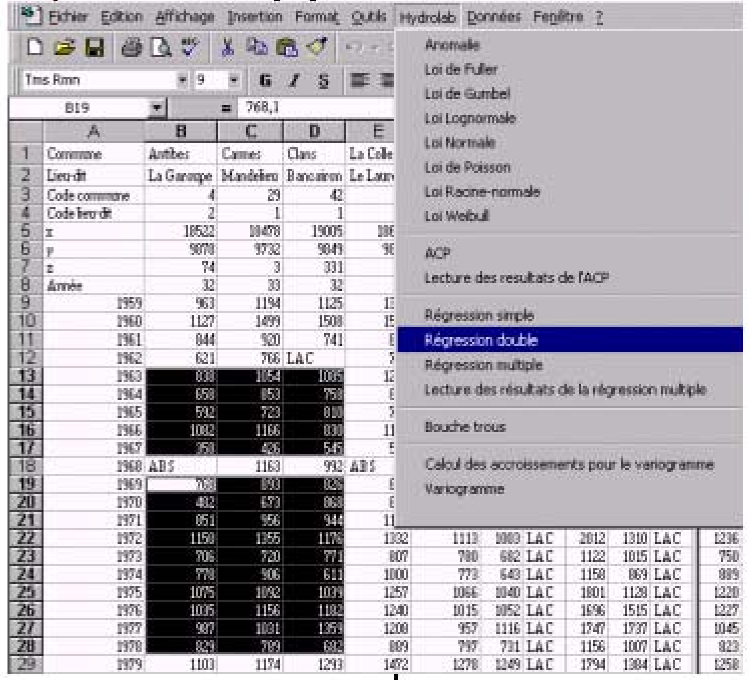
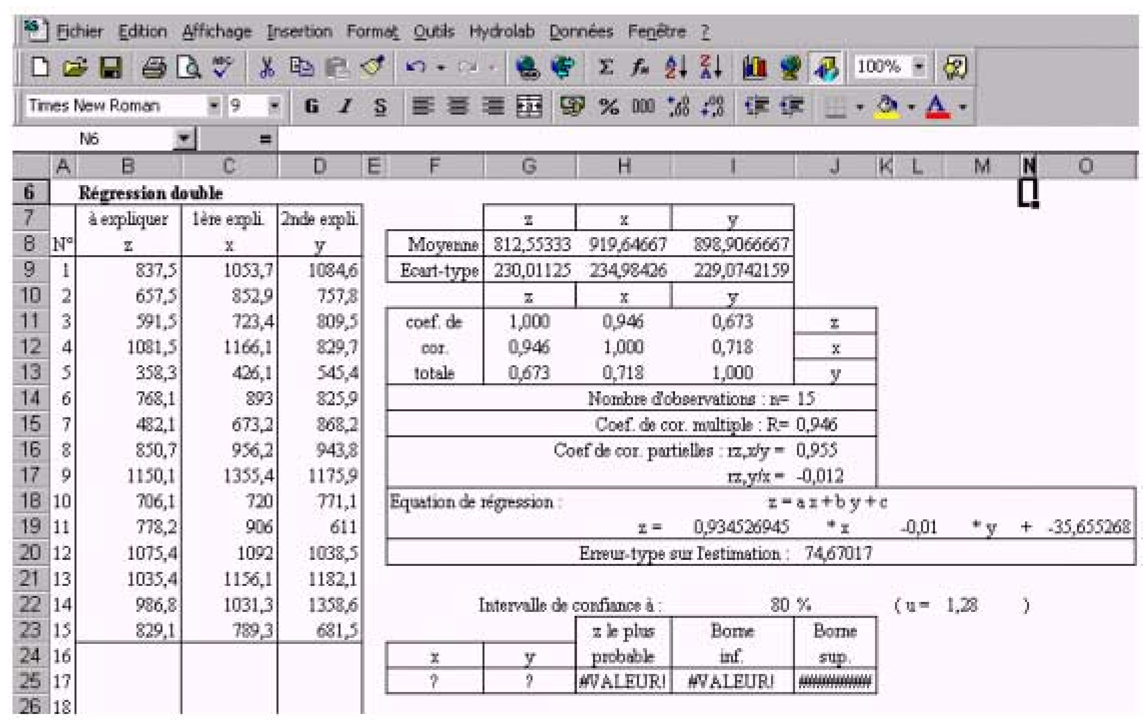


Fig.27 : Sélection des données et lancement d'une régression double

2.3.2 *Exécution de la macro*


Fichier Edition Affichage Insertion Format Outils Hydrolab Données Fenêtre ?

Times New Roman 9 — N6

Régression double

N°	à expliquer z	1ère expli. x	2nde expli. y
1	837,5	1053,7	1084,6
2	657,5	852,9	757,8
3	591,5	723,4	809,5
4	1081,5	1166,1	829,7
5	358,3	426,1	545,4
6	768,1	893	825,9
7	482,1	673,2	868,2
8	850,7	956,2	943,8
9	1150,1	1355,4	1175,9
10	706,1	720	771,1
11	778,2	906	611
12	1075,4	1092	1038,5
13	1035,4	1156,1	1182,1
14	986,8	1031,3	1358,6
15	829,1	789,3	681,5
16			
17			
18			

	z	x	y	
Moyenne	812,55333	919,64667	898,9066667	
Ecart-type	230,01125	234,98426	229,0742159	
	z	x	y	
coef. de	1,000	0,946	0,673	z
cor.	0,946	1,000	0,718	x
totale	0,673	0,718	1,000	y

Nombre d'observations : n= 15
Coef. de cor. multiple : R= 0,946
Coef de cor. partielles : rz,x/y = 0,955
rz,y/x = -0,012
Equation de régression : z = a x + b y + c
z = 0,934526945 * x -0,01 * y + -35,655268
Erreur-type sur l'estimation : 74,67017

Intervalle de confiance à : 80 % (u= 1,28)

x	y	z le plus probable	Borne inf.	Borne sup.
?	?	#VALEUR!	#VALEUR!	##########

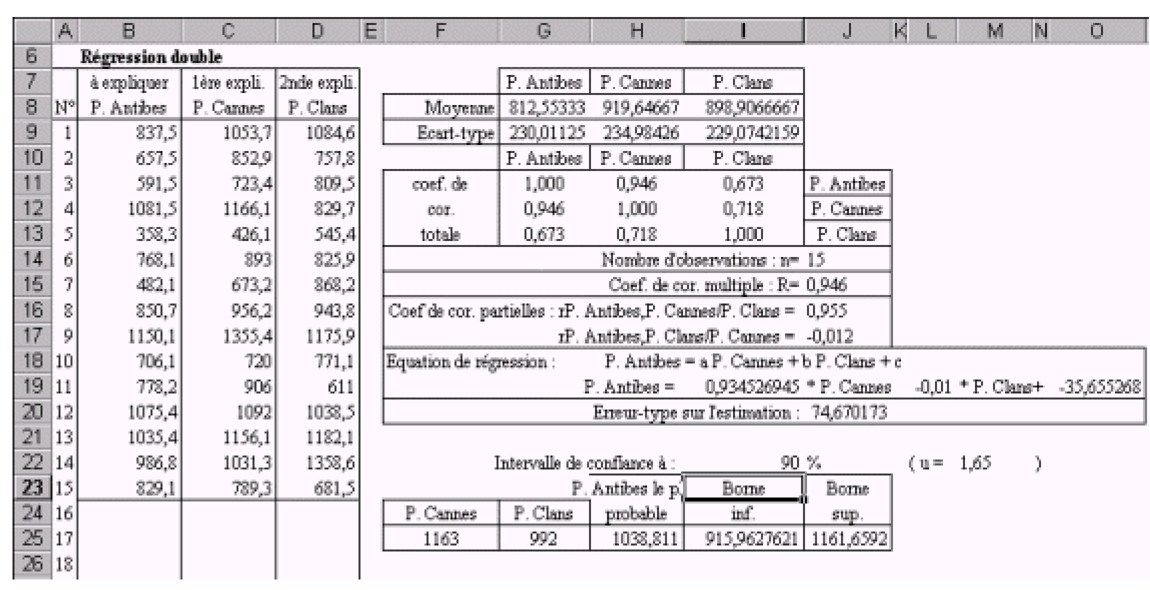
Fig. 28 : Résultats numériques d'une régression double

A cette feuille de calcul sont associées trois feuilles graphiques qui présentent les régressions partielles de la variable à expliquer par rapport aux deux variables explicatives ainsi que les reconstitutions de la variable à expliquer en regard de celles réellement observées (voir figure 30).

Pour les aspects théoriques on se reportera à ce qui est dit au paragraphe 2.6 pour ce qui concerne les régressions multiples en général.

2.3.3 *Personnalisation des résultats*

Tout comme pour les régressions simples on peut personnaliser les feuilles en tapant les noms des variables (B8 à D8), en changeant les intervalles de confiance (I22), en introduisant des valeurs des variables explicatives (F25 et G25), en ajoutant du texte sur les graphiques…


Régression double

N°	à expliquer P. Antibes	1ère expli. P. Cannes	2nde expli. P. Clans
1	837,5	1053,7	1084,6
2	657,5	852,9	757,8
3	591,5	723,4	809,5
4	1081,5	1166,1	829,7
5	358,3	426,1	545,4
6	768,1	893	825,9
7	482,1	673,2	868,2
8	850,7	956,2	943,8
9	1150,1	1355,4	1175,9
10	706,1	720	771,1
11	778,2	906	611
12	1075,4	1092	1038,5
13	1035,4	1156,1	1182,1
14	986,8	1031,3	1358,6
15	829,1	789,3	681,5
16			
17			
18			

	P. Antibes	P. Cannes	P. Clans	
Moyenne	812,55333	919,64667	898,9066667	
Ecart-type	230,01125	234,98426	229,0742159	
	P. Antibes	P. Cannes	P. Clans	
coef. de	1,000	0,946	0,673	P. Antibes
cor.	0,946	1,000	0,718	P. Cannes
totale	0,673	0,718	1,000	P. Clans

Nombre d'observations : n= 15
Coef. de cor. multiple : R= 0,946
Coef de cor. partielles : rP. Antibes,P. Cannes/P. Clans = 0,955
rP. Antibes,P. Clans/P. Cannes = -0,012
Equation de régression : P. Antibes = a P. Cannes + b P. Clans + c
P. Antibes = 0,934526945 * P. Cannes -0,01 * P. Clans+ -35,655268
Erreur-type sur l'estimation : 74,670173

Intervalle de confiance à : 90 % (u= 1,65)

P. Cannes	P. Clans	P. Antibes le plus probable	Borne inf.	Borne sup.
1163	992	1038,811	915,9627621	1161,6592

Fig. 30 : Personnalisation des résultats de la régression double

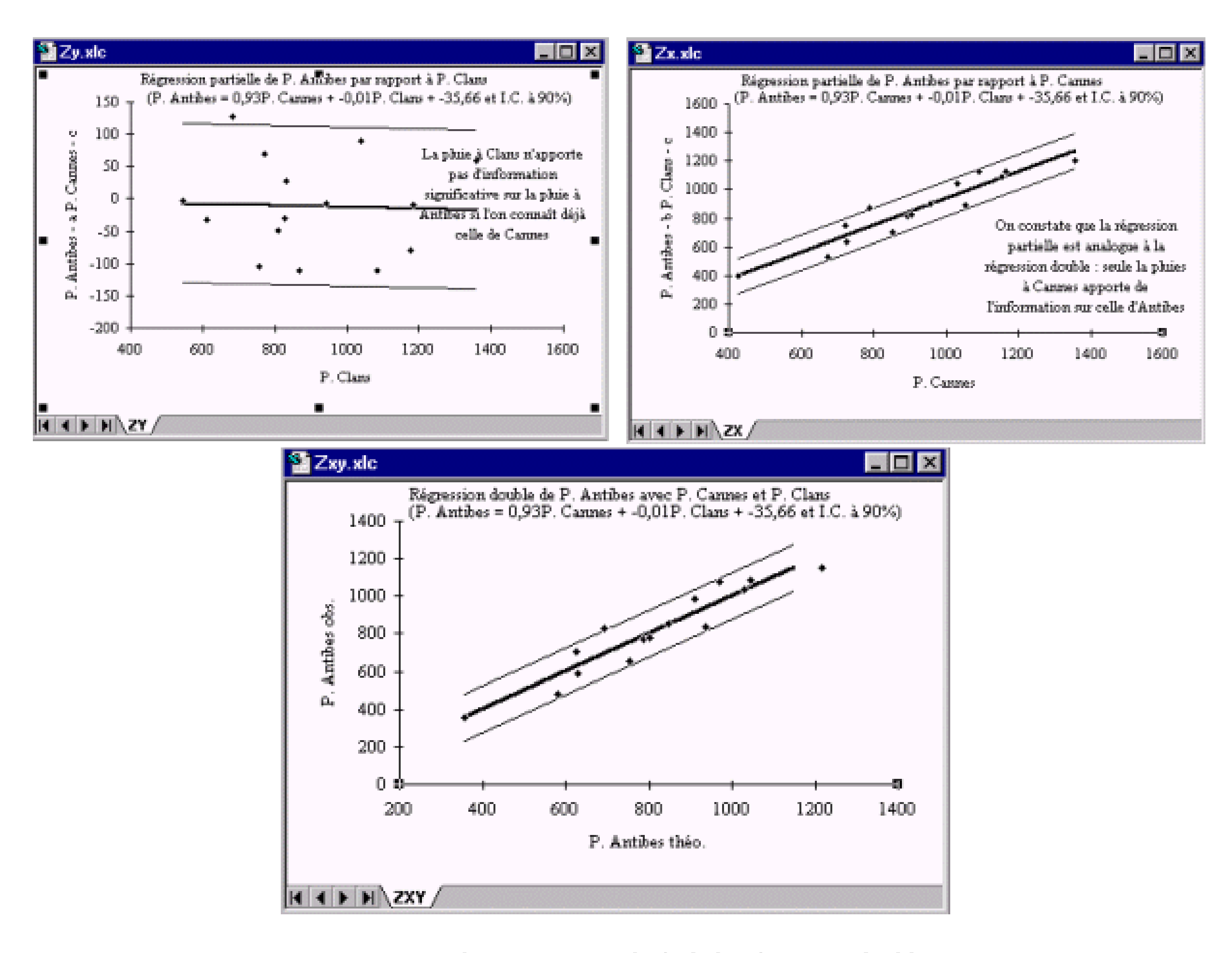


Fig. 31 : Graphiques personnalisés de la régression double

2.4 Détections d'anomalies

2.4.1 Sélection des données

La figure 32 illustre le mode de sélection des données qui est analogue à celui de la régression simple. Par contre la convention est que se trouve dans la première colonne la variable à tester et que la seconde colonne contient la variable de référence réputée sans erreurs.

Microsoft Excel - Pluies06.xls

Fichier Edition Affichage Insertion Format Outils Hydrolab Données Fenêtre ?

Tms Rmn 9

B9 = 963,4

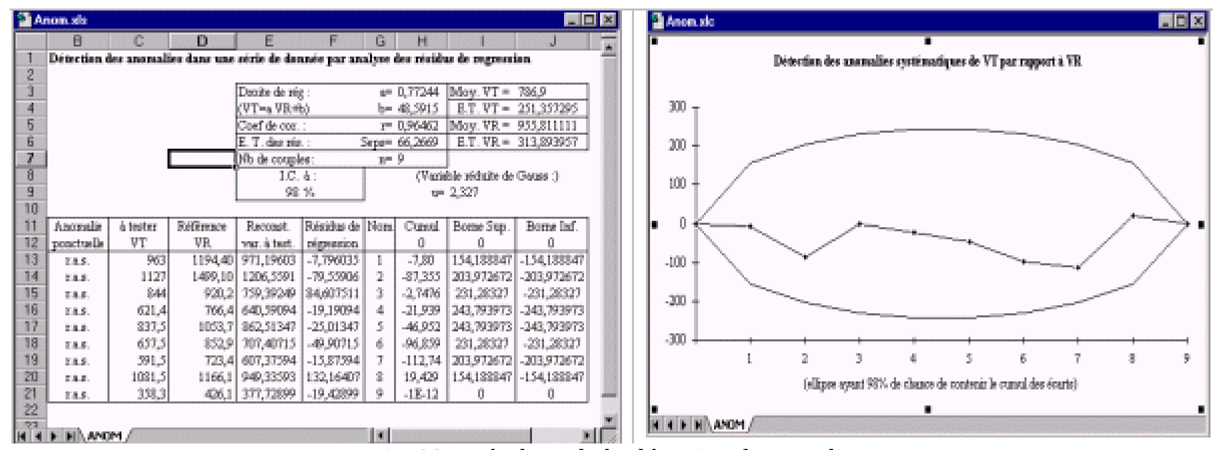
	A	B	C	D	E
1	Commune	Antibes	Cannes	Clans	La Colle
2	Lieu-dit	La Garoupe	Mandelieu	Bancairon	Le Laur
3	Code commune	4	29	42	
4	Code lieu-dit	2	1	1	
5	x	18522	18478	19005	186
6	y	9878	9732	9849	96
7	z	74	3	331	
8	Année	32	33	32	
9	1959	963	1194	1125	13
10	1960	1127	1499	1508	15
11	1961	844	920	741	6
12	1962	621	766	LAC	7
13	1963	838	1054	1085	12
14	1964	658	853	758	6
15	1965	592	723	810	7
16	1966	1082	1166	830	11
17	1967	358	426	545	5
18	1968	ABS	1163	992	ABS
19	1969	768	893	826	6
20	1970	482	673	868	6
21	1971	851	956	944	11
22	1972	1150	1355	1176	1332

Row 22 (continued): 1113 | 1003 LAC | 2012 | 1310 LAC | 1236

Hydrolab:
- Anomalie
- Loi de Fuller
- Loi de Gumbel
- Loi Lognormale
- Loi Normale
- Loi de Poisson
- Loi Racine-normale
- Loi Weibull
- ACP
- Lecture des resultats de l'ACP
- Régression simple
- Régression double
- Régression multiple
- Lecture des résultats de la régression multiple
- Bouche trous
- Calcul des accroissements pour le variogramme
- Variogramme

Fig. 32 : Sélection des séries pour une détection des anomalies

2.4.2 Exécution de la macro



Anom.xls

Détection des anomalies dans une série de donnée par analyse des résidus de regression

Droite de rég :	a= 0,77244	Moy. VT =	786,9
(VT=a VR+b)	b= 48,5915	E.T. VT =	251,357295
Coef de cor. :	r= 0,96462	Moy. VR =	955,811111
E. T. des rés. :	Sepa= 66,2669	E.T. VR =	313,893957
Nb de couples:	n= 9		

I.C. à : 98 %

(Variable réduite de Gauss :) u= 2,327

Anomalie ponctuelle	à tester VT	Référence VR	Reconst. var. à test.	Résidus de régression	Num	Cumul 0	Borne Sup. 0	Borne Inf. 0
r.a.s.	963	1194,40	971,19603	-7,796033	1	-7,80	154,188847	-154,188847
r.a.s.	1127	1499,10	1206,5591	-79,55906	2	-87,355	203,972672	-203,972672
r.a.s.	844	920,2	759,39249	84,607511	3	-2,7476	231,28327	-231,28327
r.a.s.	621,4	766,4	640,59094	-19,19094	4	-21,939	243,793973	-243,793973
r.a.s.	837,5	1053,7	862,51347	-25,01347	5	-46,952	243,793973	-243,793973
r.a.s.	657,5	852,9	707,40715	-49,90715	6	-96,859	231,28327	-231,28327
r.a.s.	591,5	723,4	607,37594	-15,87594	7	-112,74	203,972672	-203,972672
r.a.s.	1081,5	1166,1	949,33593	132,16407	8	19,429	154,188847	-154,188847
r.a.s.	358,3	426,1	377,72899	-19,42899	9	-1E-12	0	0

Fig. 33 : Résultats de la détection d'anomalies

Cette macro a pour objectif de déterminer deux types d'anomalies : les anomalies ponctuelles et les anomalies systématiques.

Dans le premier cas, on signale dans la colonne de gauche, les valeurs dont les résidus de régression dépassent une certaine probabilité d'être due au hasard.

Si comme précédemment, les variables x et y sont corrélées et que les hypothèses d'une distribution de Gauss à deux dimensions sont acceptables, on aura les relations suivantes :

$$\overline{y}_x = ax_i + b \qquad \text{et} \qquad y_i - \overline{y}_x = \varepsilon_i$$

ε_i est une variable gaussienne de moyenne nulle et d'écart-type $\sigma_y\sqrt{1-\sigma^2}$. Pour chaque valeur y_i, on calculera le ε_i correspondant et sa fréquence théorique. Les valeurs de ε_i ayant des fréquences très rares correspondent à des y_i douteux. On peut, par cette méthode, détecter des erreurs accidentelles qui n'apparaissent pas à l'étude des distributions marginales. Dans l'exemple de la figure 33 aucune valeur ne paraît anormale au seuil standard de 98%.

Les anomalies systématiques sont détectées sur la base de l'analyse du cumul des résidus de régression : Comme on vient de le dire au paragraphe précédent, le résidu ε_i est une variable aléatoire gaussienne de moyenne nulle et d'écart-type $\sigma_y\sqrt{1-\sigma^2}$. On définit alors la variable S_i, cumul des i premiers résidus :

$$Z_i = \sum_{j=1}^{i} \varepsilon_j$$

Cette variable S_i est une variable aléatoire de moyenne nulle et d'écart-type : $\sigma_{Z_i} = \sigma_\varepsilon \sqrt{\frac{i(n-1)}{n-1}}$

Si on se fixe par exemple un intervalle de confiance à 98%, il y à deux chances sur cent pour que S_i soit extérieur au segment :

$$\left[-u_{0.99}\sigma_{Z_i}, +u_{0.99}\sigma_{Z_i}\right] \qquad \text{Soit encore :} \qquad \left[-2.33\sigma_{Z_i}, +2.33\sigma_{Z_i}\right]$$

Lorsque j varie, le lieu des limites des segments ayant la même probabilité de contenir sont des ellipses passant par Z=0 pour i=0 et Z=0 pour i=n.

Sur le graphique ci-dessous, on a porté les cumuls des résidus en fonction de i et la courbe restant à l'intérieur de l'ellipse à 98%, les deux séries paraissent homogènes.

2.4.3 Personnalisation des résultats

Anom.xls

Détection des anomalies dans une série de donnée par analyse des résidus de regression

Droite de rég :	a= 0,77244
(Antibes=a Cannes+b)	b= 48,5915
Coef de cor. :	r= 0,96462
E. T. des rés. :	Sepe= 66,2669
Nb de couples :	n= 9
I.C. à : 95 %	(Variable réduite de Gauss :) u= 1,960

Moy. Antibes =	786,9
E.T. Antibes =	251,357295
Moy. Cannes =	955,811111
E.T. Cannes =	313,893957

Anomalie ponctuelle	à tester Antibes	Référence Cannes	Reconst. var. à test.	Résidus de régression	Nom	Cumul 0	Borne Sup. 0	Borne Inf. 0
r.a.s.	963	1194,40	971,19603	-7,796035	1959	-7,80	129,9092906	-129,909291
r.a.s.	1127	1499,10	1206,5591	-79,55906	1960	-87,355	171,8538379	-171,853838
r.a.s.	844	920,2	759,39249	84,607511	1961	-2,7476	194,8639358	-194,863936
r.a.s.	621,4	766,4	640,59094	-19,19094	1962	-21,939	205,4046237	-205,404624
r.a.s.	837,5	1053,7	862,51347	-25,01347	1963	-46,952	205,4046237	-205,404624
r.a.s.	657,5	852,9	707,40715	-49,90715	1964	-96,859	194,8639358	-194,863936
r.a.s.	591,5	723,4	607,37594	-15,87594	1965	-112,74	171,8538379	-171,853838
à vérifier	1081,5	1166,1	949,33593	132,16407	1966	19,429	129,9092906	-129,909291
r.a.s.	358,3	426,1	377,72899	-19,42899	1967	-1E-12	0	0

ANOM

Fig. 34 : Détection d'une anomalie ponctuelle.

Sur la figure 34 on voit que après avoir collé les dates à partir de la cellule G12, indiqué les noms des variables (en C12 et D12) et modifié l'intervalle de confiance en le passant à 95% (cellule E9), la méthode détecte une anomalie en 1966 où la pluie à Antibes est de 1081.5 mm alors que compte tenu des 1166 mm à Cannes on aurait pu s'attendre à 949 mm. Cet écart de 132 mm a moins de 5% de chance d'être du au hasard et la macro nous indique que cette valeur est à vérifier.

Sur la figure 35 nous avons traité toutes les pluies communes aux stations d'Antibes et Cannes (1959 - 1992 et à l'exception des deux années 68 et 87) ; on remarque très nettement une rupture de stationnarité de par et d'autre des années 70 à 72.

En fait, il s'agit ici d'un déplacement de la station de Cannes qui n'est donc pas la référence.

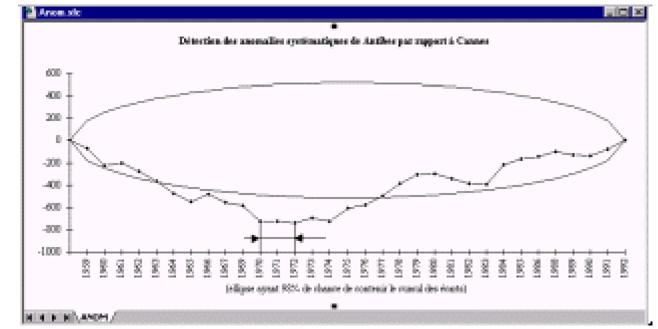


Fig. 35 : Détection d'une rupture de stationnarité.

2.5 Analyse en composantes principales

2.5.1 Préparation des données

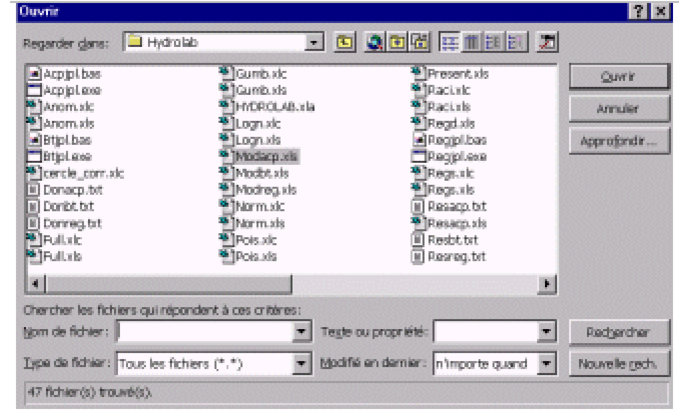


Fig. 36 : Ouverture de la feuille modèle pour ACP

La nature des données étant beaucoup plus complexe nous avons préféré passer par l'intermédiaire d'une feuille de calcul spécifique notée "Modacp.xls". Cette feuille est un document modèle dans lequel vous aller pouvoir regrouper tranquillement par le biais de "Copier - Coller" les informations à traiter.

La première étape consiste à ouvrir dans le répertoire HYDROLAB le document "Modacp.xls" comme indiqué à la figure 36

Cette feuille comporte un certain nombre de cellules prévues pour l'introduction des données. La ligne six, à partir de la colonne B contiendra le nom des variables (des alphanumériques ou des nombres) de même la colonne A à partir de la ligne sept contiendra le nom des observations (des alphanumériques ou des nombres). La plage des cellules à gauche et en dessous de la cellule B7 est destinée à recevoir les données numériques (il ne faut absolument pas qu'il y ait de lacunes: blancs ou alphanumériques).

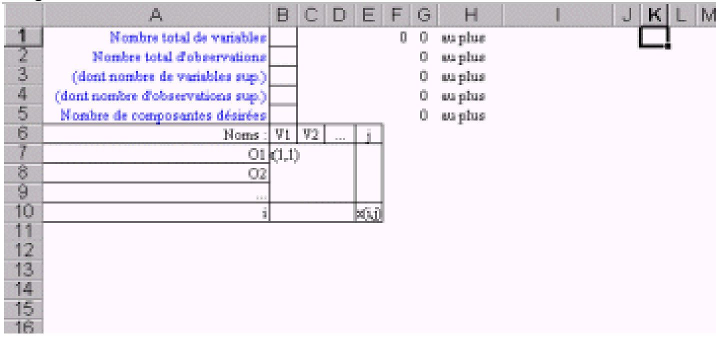


Fig. : 37 Aspect de la feuille "Modacp.xls" à l'ouverture

	A	B	C	D	E	F	G	H	I	J	K
1	Commune	Antibes	Cannes	Clans	La Colle	Colomars	Contes	Contes	Coursegoules	Entraunes	Entraun
2	Lieu-dit	La Garoupe	Mandelieu	Bancairon	Le Lauron	Cie des Eau	La Grave	Sclos	Place Neuve	Le Clos	Esteinc
3	Code commune	4	29	42	44	46	48	48	50	56	56
4	Code lieu-dit	2	1	1	1	1	1		1	1	3
5	x	18522	18478	19005	18668	18750	18792		18770	19190	19264
6	y	9878	9732	9849	9815	9931	10010		9788	9529	9526
7	z	74	3	331	44	334	188		1000	1250	1850
8	Année	32	33	32	32	30	27	8	33	27	13
9	1959	963	1194	1125	1300	1383	LAC	1247	2034	LAC	1547
10	1960	1127	1499	1508	1535	1469	LAC	1383	2324	LAC	1890
11	1961	844	920	741	884	921	LAC	940	1165	LAC	1039
12	1962	621	766	LAC	771	788	LAC	752	1193	LAC	1092
13	1963	838	1054	1085	1207	1146	LAC	1055	1668	LAC	1590
14	1964	658	853	758	892	705	772	LAC	1320	1014	1108
15	1965	592	723	810	760	817	878	LAC	1062	1016	1050
16	1966	1082	1166	830	1190	1009	924	894	1276	961	1132
17	1967	358	426	545	542	555	525	564	748	714	743
18	1968	ABS	1163	992	ABS	ABS	ABS	975	2265	ABS	ABS
19	1969	768	893	826	883	786	711	ABS	1347	1093	LAC
20	1970	482	673	868	890	638	766	LAC	1232	1095	LAC
21	1971	851	956	944	1109	1059	931	LAC	1468	1061	LAC
22	1972	1150	1355	1176	1332	1113	1003	LAC	2012	1310	LAC
23	1973	706	720	771	807	780	682	LAC	1122	1015	LAC
24	1974	778	906	611	1000	773	643	LAC	1158	869	LAC
25	1975	1075	1092	1039	1257	1066	1040	LAC	1801	1128	LAC
26	1976	1035	1156	1182	1240	1015	1052	LAC	1696	1515	LAC
27	1977	987	1031	1359	1208	957	1116	LAC	1747	1737	LAC
28	1978	829	789	682	889	797	731	LAC	1156	1007	LAC
29	1979	1103	1174	1293	1472	1278	1249	LAC	1794	1384	LAC
30	1980	631	667	585	674	726	677	LAC	937	675	LAC
31	1981	620	730	948	879	801	857	LAC	1178	1114	LAC

Fig. : 38 Sélection des données pour l'ACP

La figure 38 donne un exemple de sélection des données à traiter par ACP, ici, on a sélectionné 8 variables (les postes pluviométriques et 13 observations (les pluies annuelles de 69 à 81). Une fois cette plage de cellules sélectionnée, on vient de la copier puis de la coller dans la feuille "Modacp.xls" comme l'illustre la figure 39.

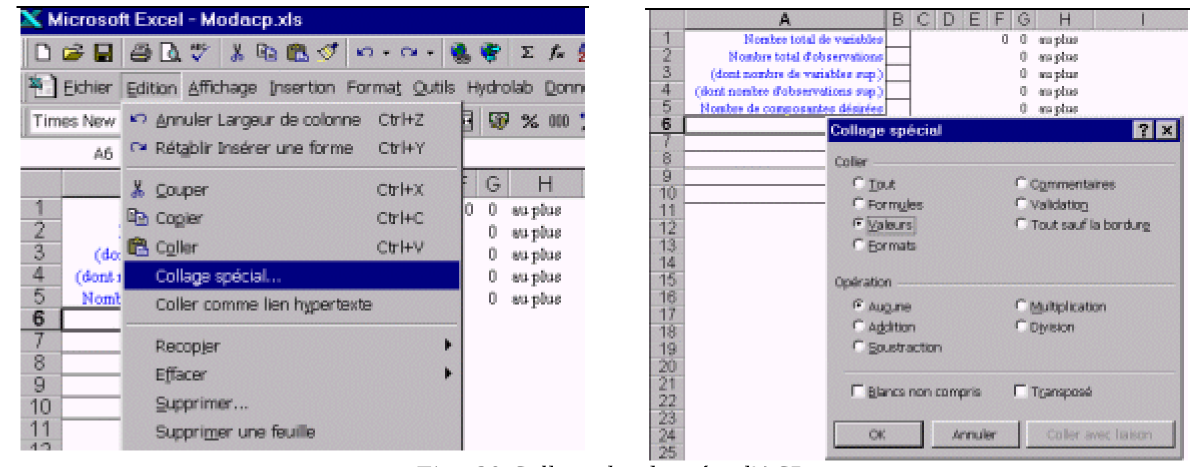


Fig. :39 Collage des données d'ACP

Il est de loin préferable de procéder par un collage spécial des valeurs, en effet certaines de vos données de départ peuvent être des résultats de formules. Dans l'exemple ci-dessus le transfert a été fait en une seule fois (nom des variables, nom des observations et les valeurs numériques), mais vous pouvez procéder en plusieurs étapes si cela vous est plus commode. Ces différents collages effectués vous devriez obtenir quelque chose d'analogue à la figure 40.

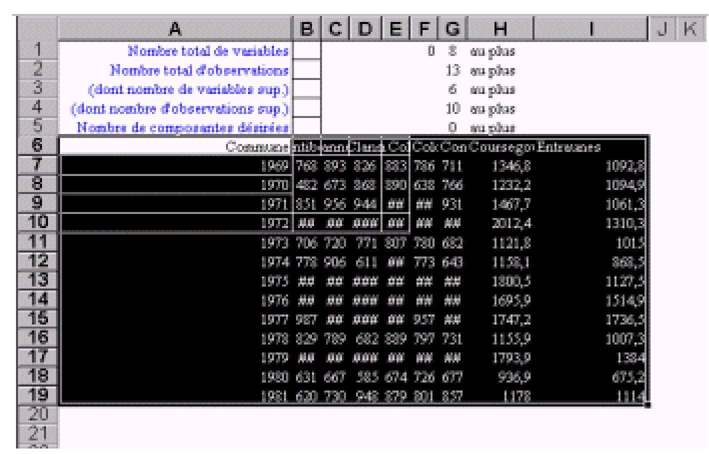


Fig. : 40 Introduction des paramètres de l'ACP

Ne vous inquiétez pas du format des colonnes (des nombres remplacés par des #, des noms de variables illisibles), qui sera géré automatiquement. Par contre il vous reste à saisir :

- le nombre total de variables à traiter, (dans la feuille on vous indique la valeur maximale compte tenu des données introduites, mais vous pourriez ne vouloir traiter que les x<8 premières).
- Le nombre total d'observations (dans la feuille on vous indique la valeur maximale compte tenu des données introduites, mais vous pourriez ne vouloir traiter que les y<13 premières).
- Le nombre de variables supplémentaires (ce sont des variables qui n'entreront pas dans le calcul des composantes mais qui seront néanmoins projetés dessus). Ces variables supplémentaires correspondent aux colonnes les plus à droite.
- Le nombre d'observations supplémentaires (ce sont des observations qui n'entreront pas dans le calcul de la matrice de coefficients de corrélation mais qui seront néanmoins projetés sur les composantes). Ce seront les observations figurant dans les dernières lignes.
- Enfin, indiquez le nombre de composantes désirées ; ce nombre doit être notablement inférieur au nombre de variables entrant dans les calculs (nombre total des variables moins nombre de variables supplémentaires).

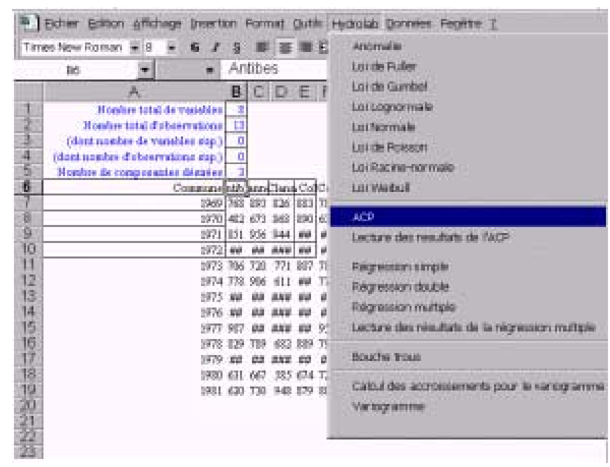


Fig. :41 Lancement de l'ACP

2.5.2 *Exécution de l'ACP*

Il suffit alors de lancer la commande ACP comme l'indique la figure 41.

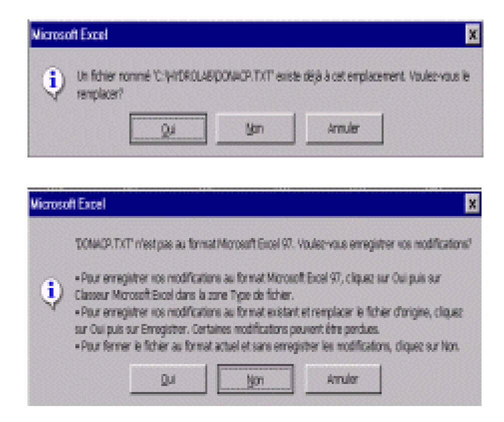


Fig. :42 Boites de dialogue de l'ACP

A la première boite de dialogue il convient de répondre oui. (la macro utilise un fichier temporaire déjà existant "Donacp.txt".

Pour la seconde boite de dialogue il convient de répondre non, en effet, HYDROLAB est conçu pour tourner indifféremment dans les versions 4, 5 et 97 d'Excel, les fichiers sont donc tous à sauvegarder en version 4 (compatibilité ascendante).

Vous verrez apparaître alors la fenêtre DOS du module exécutable. Cette fenêtre permet de suivre sommairement le déroulement des opérations. Il se peut que l'on reste bloqué dans cette fenêtre si l'on a commis des erreurs dans la construction du tableau de données (Modacp.xls). Les erreurs les plus banales sont un tableau comportant des valeurs non numériques et des variables étant des combinaisons linéaires des autres. En cas de blocage, revenir à la saisie des données.

Suivant les versions, cette fenêtre se ferme automatiquement ou sinon il vous suffit de cliquer sur la croix en haut à droite de la fenêtre (figure 43)



Fig. : 43 Fenêtre du module exécutable de l'ACP

2.5.3 Visualisation des résultats de l'ACP

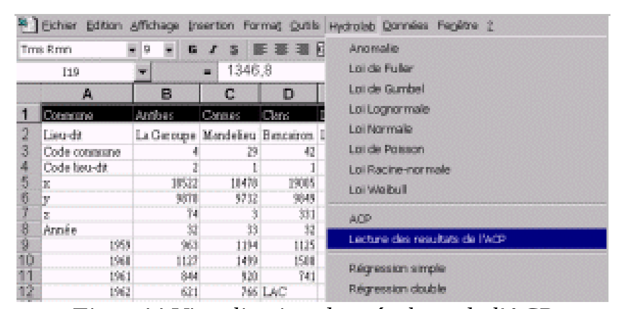


Fig. : 44 Visualisation des résultats de l'ACP

Il vous suffit alors de consulter les résultats. Pendant quelques instant la commande va récupérer les résultats du module exécutable et les remettre sous la forme d'un tableau Excel nommé "Resacp.xls". Ce fichier sera écrasé à chaque nouvelle utilisation. Une boite de dialogue vous demandera l'autorisation de remplacer l'ancienne version, il convient de répondre par oui.

	A	B	C	D	E	F	G	H	I
1	es de l'Analyse en Composantes Principales :								
2	8	variables actives							
3	0	variables supplementaires							
4	13	observations actives							
5	0	observations supplementaires							
6	3	premieres composantes principales etudiees							
7	Moyennes des observations actives :								
8		847,31	934,00	944,92	1049,23	906,85	881,38	1434,46	1154,08
9	Ecarts-types des observations actives :								
10		209,54	217,81	253,57	237,45	187,04	196,98	338,85	276,46
11	Matrice des coefficients de correlation :								
12		Antibes	Cannes	Clans	La Colle	Colomar	Contes	Courseg	Entraun
13	Antibes	1,00	0,95	0,71	0,89	0,90	0,78	0,89	0,62
14	Cannes	0,95	1,00	0,74	0,92	0,87	0,77	0,94	0,64
15	Clans	0,71	0,74	1,00	0,84	0,76	0,94	0,88	0,93
16	La Colle	0,89	0,92	0,84	1,00	0,92	0,90	0,94	0,72
17	Colomars	0,90	0,87	0,76	0,92	1,00	0,88	0,85	0,57
18	Contes	0,78	0,77	0,94	0,90	0,88	1,00	0,87	0,81
19	Coursegoules	0,89	0,94	0,88	0,94	0,85	0,87	1,00	0,77
20	Entraunes	0,62	0,64	0,93	0,72	0,57	0,81	0,77	1,00
21	Valeurs propres :								
22		C1	C 2	C 3					
23		6,82	0,71	0,26					
24	Variances expliquees :								
25		C1	C 2	C 3					
26		85%	9%	3%					
27	Variances cumulees :								
28		C1	C 2	C 3					
29		85%	94%	97%					

	A	B	C	D	E
30	Projections des variables actives :				
31		C1	C 2	C 3	
32	Antibes	0,91	-0,32	0,13	
33	Cannes	0,93	-0,28	0,21	
34	Clans	0,92	0,37	-0,05	
35	La Colle	0,97	-0,11	-0,05	
36	Colomars	0,92	-0,27	-0,26	
37	Contes	0,94	0,17	-0,27	
38	Coursegoules	0,97	-0,04	0,14	
39	Entraunes	0,82	0,53	0,18	
40	Projections des variables supplementaires :				
41		C1	C 2	C 3	
42	Projections des observations actives				
43		C1	C 2	C 3	
44	1969,00	-0,51	-0,01	1,15	
45	1970,00	-0,92	1,47	-0,05	
46	1971,00	0,17	-0,59	-1,22	
47	1972,00	1,29	-1,05	1,39	
48	1973,00	-0,88	0,18	0,11	
49	1974,00	-0,77	-1,20	0,88	
50	1975,00	0,79	-0,97	-0,32	
51	1976,00	0,96	0,48	0,75	
52	1977,00	1,05	1,97	0,66	
53	1978,00	-0,70	-0,45	0,35	
54	1979,00	1,53	-0,28	-1,93	
55	1980,00	-1,41	-0,64	-0,82	
56	1981,00	-0,59	1,09	-0,95	
57	ojections des observations supplementaires :				
58		C1	C 2	C 3	

Fig. : 45 Exemple de résultats d'une analyse en composantes principales

La programmation de l'analyse en composantes principales s'inspire des ouvrages de LEBART & al. (1979) et de LEGRAS (1972). Ce programme a été écrit en Basic compilé, seul le module exécutable est indispensable mais vous pouvez consulter la source "Acpjpl.bas" dans le répertoire HYDROLAB.

Il faut noter que nous avons choisi de présenter les projections des observations sur les composantes de façon centrée réduite. Dans certains autres programmes ces projections sont données avec une variance égale à la valeur propre correspondante. Il ne faut donc pas s'alarmer si vous constatez des écarts avec les résultats d'autres programmes ; la convention n'est pas la même et il suffit d'un terme multiplicatif pour homogénéiser les résultats.

Si vous voulez conserver ces résultats il suffit de les enregistrer sous un autre nom.

2.6 Régressions multiples

2.6.1 Préparation des données

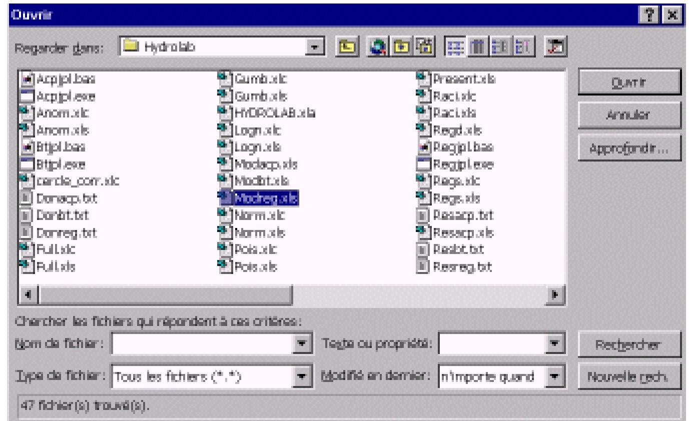


Fig. 46 : Ouverture de la feuille modèle pour régression multiple

La préparation des données est tout à fait analogue à ce que nous avons décrit pour l'analyse en composantes principales. Nous passons par l'intermédiaire d'une feuille de calcul spécifique notée "Modreg.xls". Cette feuille est un document modèle dans lequel vous aller pouvoir regrouper tranquillement par le biais de "Copier - Coller" les informations à traiter.

La première étape consiste à ouvrir dans le répertoire HYDROLAB le document "Modreg.xls" comme indiqué à la figure 46

Cette feuille comporte un certain nombre de cellules prévues pour l'introduction des données. La ligne trois, à partir de la colonne B contiendra le nom des variables (des alphanumériques ou des nombres) de même la colonne A à partir de la ligne quatre contiendra le nom des observations (des alphanumériques ou des nombres). La plage des cellules à gauche et en dessous de la cellule B4 est destinée à recevoir les données numériques (il ne faut absolument pas qu'il y ait de lacunes: blancs ou alphanumériques). Il est indispensable de mettre dans la première colonne la variable à expliquer, puis les variables explicatives dans les colonnes suivantes.

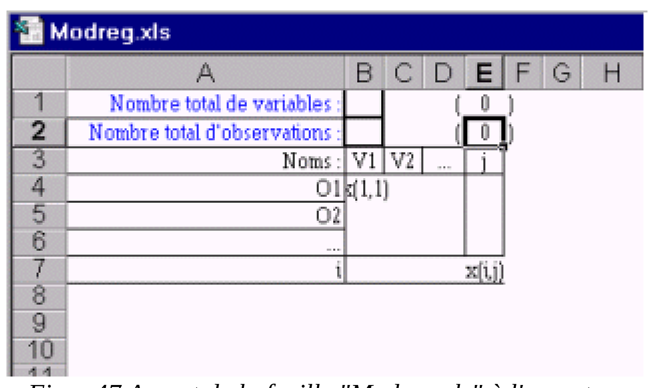


Fig. : 47 Aspect de la feuille "Modreg.xls" à l'ouverture

	A	B	C	D	E	F	G	H	I	J	K
1	Commune	Coursegoules	Antibes	Cannes	Clans	La Colle	Colomars	Contes	Contes	Entrau	Entraur
2	Lieu-dit	Place Neuve	La Garoupe	Mandelieu	Bancairon	Le Lauron	Cie des Eau	La Gra	Sclos d	Le Clo	Esteinc
3	Code commune	50	4	29	42	44	46	48	48	56	56
4	Code lieu-dit	1	2	1	1	1	1	1		1	3
5	x	18770	18522	18478	19005	18668	18750	18792		19190	19264
6	y	9788	9878	9732	9849	9815	9931	10010		9529	9526
7	z	1000	74	3	331	44	334	188		1250	1850
8	Année	33	32	33	32	32	30	27	8	27	13
9	1959	2034	963	1194	1125	1300	1383	LAC	1247	LAC	1547
10	1960	2324	1127	1499	1508	1535	1469	LAC	1383	LAC	1890
11	1961	1165	844	920	741	884	921	LAC	948	LAC	1039
12	1962	1193	621	766	LAC	771	788	LAC	752	LAC	1092
13	1963	1668	838	1054	1085	1207	1146	LAC	1055	LAC	1590
14	1964	1320	658	853	758	892	705	772	LAC	1014	1108
15	1965	1062	592	723	810	760	817	878	LAC	1016	1050
16	1966	1276	1082	1166	830	1190	1009	924	894	961	1132
17	1967	748	358	426	545	542	555	525	564	714	743
18	1968	2265	ABS	1163	992	ABS	ABS	ABS	975	ABS	ABS
19	1969	1347	768	893	826	883	706	711	ABS	1093	LAC
20	1970	1232	482	673	868	890	638	766	LAC	1095	LAC
21	1971	1468	851	956	944	1109	1059	931	LAC	1061	LAC
22	1972	2012	1150	1355	1176	1332	1113	1003	LAC	1310	LAC
23	1973	1122	706	720	771	807	780	682	LAC	1015	LAC
24	1974	1150	778	906	611	1000	773	643	LAC	869	LAC
25	1975	1801	1075	1092	1039	1257	1066	1040	LAC	1128	LAC
26	1976	1696	1035	1156	1182	1240	1015	1052	LAC	1515	LAC
27	1977	1747	987	1031	1359	1208	957	1116	LAC	1737	LAC
28	1978	1156	829	789	682	889	797	731	LAC	1007	LAC
29	1979	1794	1103	1174	1293	1472	1278	1249	LAC	1384	LAC
30	1980	937	631	667	585	674	726	677	LAC	675	LAC
31	1981	1178	620	730	948	879	801	857	LAC	1114	LAC
32	1982	949	571	652	710	689	606	614	LAC	950	LAC

Fig. : 48 Sélection des données pour la régression multiple

La figure 48 donne un exemple de sélection des données où on cherche à expliquer les pluies à Coursegoule (la colonne correspondante de la feuille de données "pluies06.xls" a été permutée) à partir de sept autres stations du département. La régression est étudiée à partir de 13 observations (les pluies annuelles de 69 à 81). Une fois cette plage de cellules sélectionnée, on vient de la copier puis de la coller dans la feuille "Modreg.xls" comme l'illustre la figure 49.

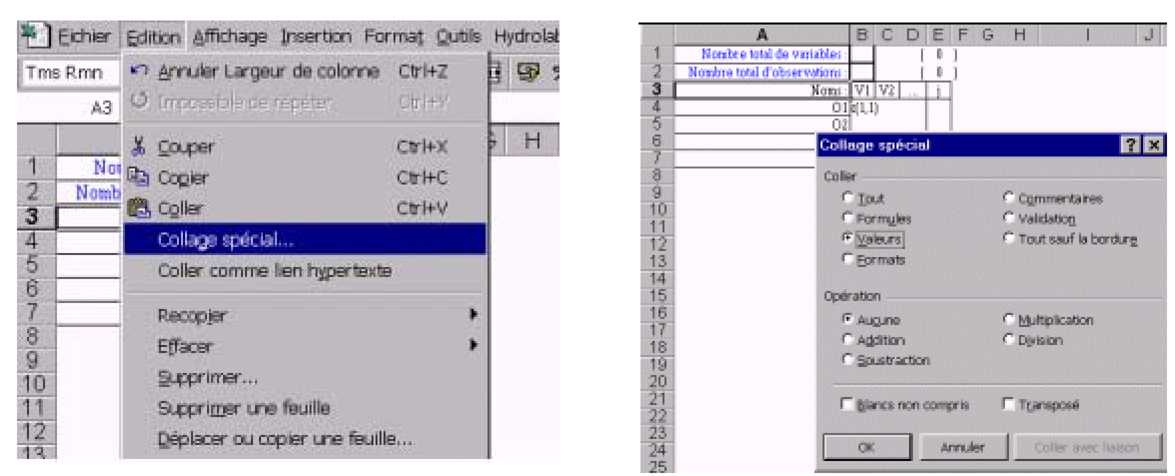


Fig. :49 Collage des données de la régression multiple

Il est de loin préférable de procéder par un collage spécial des valeurs, en effet certaines de vos données de départ peuvent être des résultats de formules. Dans l'exemple ci-dessus le transfert a été fait en une seule fois (nom des variables, nom des observations et les valeurs numériques), mais vous pouvez procéder en plusieurs étapes si cela vous est plus commode. Ces différents collages effectués vous devriez obtenir quelque chose d'analogue à la figure 50.

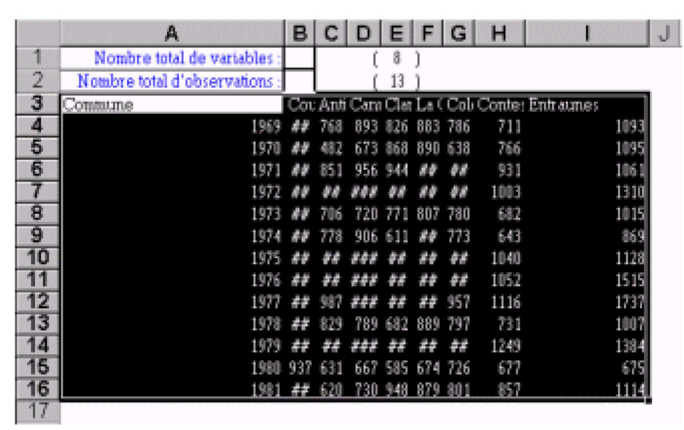


Fig. : 50 Introduction des paramètres de la régression multiple

Ne vous inquiétez pas du format des colonnes (des nombres remplacés par des #, des noms de variables illisibles), qui sera géré automatiquement. Par contre il vous reste à saisir :

- le nombre total de variables, (le nombre de variables explicatives plus un). Si vous tapez un nombre n inférieur à celui proposé (par exemple six au lieu de huit), la première variable sera expliquée par les n-1 variables suivantes.

- Le nombre total d'observations (dans la feuille on vous indique la valeur maximale compte tenu des données introduites, mais vous pourriez ne vouloir traiter que les y<13 premières).

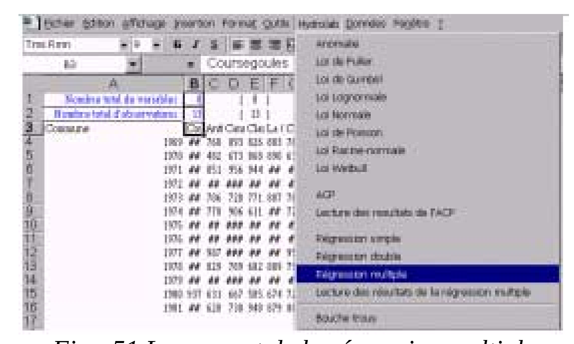


Fig. :51 Lancement de la régression multiple

2.6.2 *Exécution de la régression multiple*

Il suffit alors de lancer la commande Régression Multiple comme l'indique la figure 51.

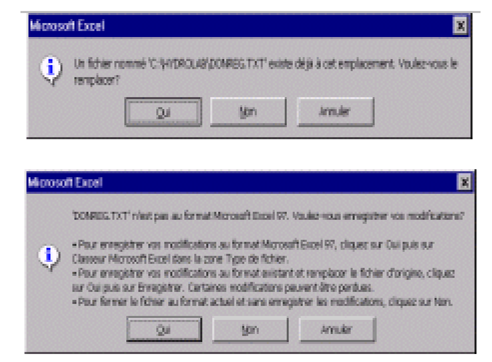


Fig. :52 Boites de dialogue de la régression multiple

A la première boite de dialogue il convient de répondre oui. (la macro utilise un fichier temporaire déjà existant "Donreg.txt".

Pour la seconde boite de dialogue il convient de répondre non, en effet, HYDROLAB est conçu pour tourner indifféremment dans les versions 4, 5 et 97 d'Excel, les fichiers sont donc tous à sauvegarder en version 4 (compatibilité ascendante).

Vous verrez apparaître alors la fenêtre DOS du module exécutable. Cette fenêtre permet de suivre sommairement le déroulement des opérations. Il se peut que l'on reste bloqué dans cette fenêtre si l'on a commis des erreurs dans la construction du tableau de données (Modreg.xls). Les erreurs les plus banales sont un tableau comportant des valeurs non numériques et des variables étant des combinaisons linéaires des autres. En cas de blocage, revenir à la saisie des données.

Suivant les versions, cette fenêtre se ferme automatiquement ou sinon il vous suffit de cliquer sur la croix en haut à droite de la fenêtre (figure 53)



Fig. : 53 Fenêtre du module exécutable de la régression multiple

2.6.3 *Visualisation des résultats de la régression multiple*

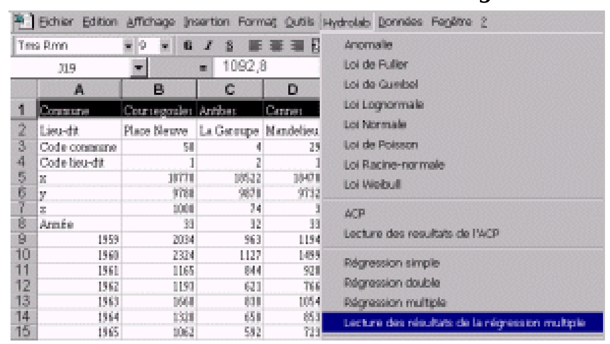


Fig. : 54 Visualisation des résultats de la régression multiple

Il vous suffit alors de consulter les résultats. Pendant quelques instant la commande va récupérer les résultats du module exécutable et les remettre sous la forme d'un tableau Excel nommé "Resreg.xls". Ce fichier sera écrasé à chaque nouvelle utilisation. Une boite de dialogue vous demandera l'autorisation de remplacer l'ancienne version, il convient de répondre par oui.

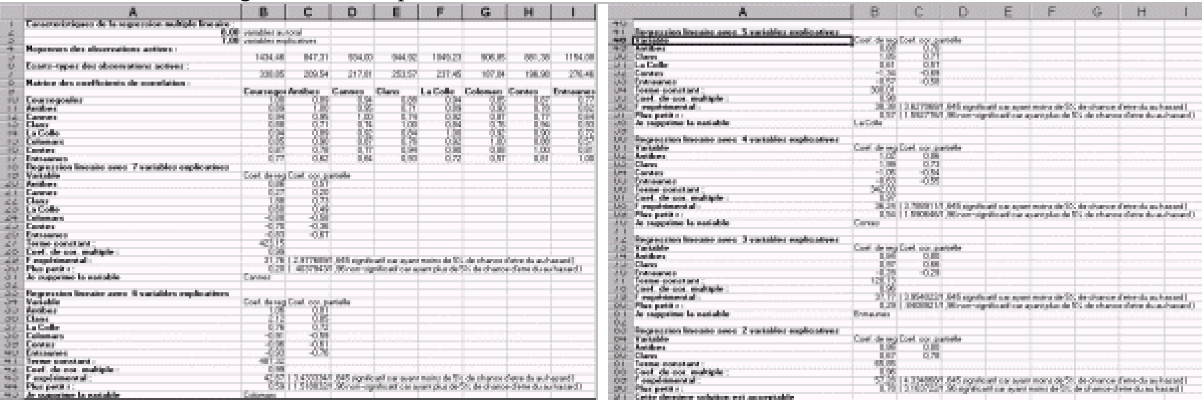


Fig. : 55 Exemple de résultats d'une régression multiple

La programmation de la régression multiple s'inspire des travaux de TOUCHEBOEUF de LUSIGNY . Ce programme a été écrit en Basic compilé, seul le module exécutable est indispensable mais vous pouvez consulter la source "Regjpl.bas" dans le répertoire HYDROLAB.

Nous avons adopté la démarche dite régressive (backward elimination), elle consiste dans un premier temps à faire une régression multiple sur l'ensemble des variables explicatives envisagées. Une fois ce premier résultat établi, on vérifie la signification des coefficients de corrélation partielle (en fait, il suffit de tester le plus petit coefficient de corrélation partielle en valeur absolue). On teste cette signification par le test de Student au seuil de 5% (seuil choisi à priori est non modifiable). Ensuite on teste le coefficient de corrélation multiple par le test de Fisher-Snedecor au seuil de 5%. Si l'un de ces deux tests est négatif on supprime la variable explicative correspondant au plus petit coefficient de corrélation partielle, et on recommence avec une variable explicative en moins. Le programme s'arrête lorsque tous les tests sont positifs.

Sur la figure 55, vous trouvez tous les résultats intermédiaires justifiant la démarche. Successivement les variables explicatives Cannes, Colomar, la Colle, Contes et enfin Entraunes sont éliminées. Dans ce cas de figure, la régression finale revient à expliquer les pluies à Coursegoule par les seules observations d'Antibes et de Clans. La régression à pour expression :

$$P_{Coursegoules} = 0.86 P_{Antibes} + 0.67 P_{Clans} + 65.85$$

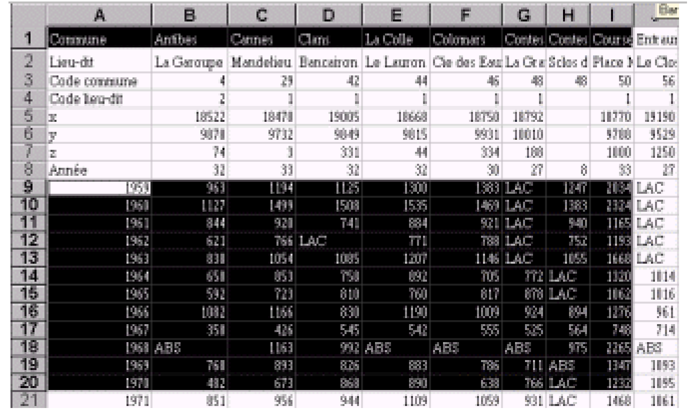
Le coefficient de corrélation multiple est de 0.96 et les coefficients de corrélation partielle avec Antibes et Clans sont respectivement 0.80 et 0.78.

Si vous voulez conserver ces résultats il suffit de les enregistrer sous un autre nom.

2.7 Comblement de lacunes (Bouche trous)

L'objectif de cette commande est de combler des lacunes dans une matrice de données. Les variables sont comme toujours disposées suivant des colonnes et les observations suivant des lignes. Dans cette matrice de données certaines valeurs absentes seront remplacées par un texte contenant la lettre a (par exemple : "Abs.", "Lacune", …). On trouvera plus loin la justification du calcul il suffit pour l'instant de savoir que le nombre de composantes principales à prendre en compte est au libre choix de l'utilisateur. Pour des données pluviométriques par exemple, il est recommandé de ne retenir que les composantes présentants une structure spatiale. L'opérateur a également le choix du nombre d'itérations à effectuer (10 à 15 sont généralement suffisantes). Pour plus de détail sur cette méthode on peut se référer à l'extrait du cours LABORDE (1998) donné en annexe I.

2.7.1 *Préparation des données*



	A	B	C	D	E	F	G	H	I	J
1	Commune	Antibes	Cannes	Clans	La Colle	Colomars	Contes	Contes	Course	Entraun
2	Lieu-dit	La Garoupe	Mandelieu	Bancairon	Le Lauron	Cie des Eau	La Gra	Sclos d	Place M	Le Clo
3	Code commune	4	29	42	44	46	48	48	50	56
4	Code lieu-dit	2	1	1	1	1	1		1	1
5	x	18522	18478	19005	18668	18750	18792		18770	19190
6	y	9878	9732	9849	9815	9931	10010		9708	9529
7	z	74	3	331	44	334	188		1000	1250
8	Année	32	33	32	32	30	27	8	33	27
9	1959	963	1194	1125	1300	1383	LAC	1247	2034	LAC
10	1960	1127	1499	1508	1535	1469	LAC	1383	2324	LAC
11	1961	844	920	741	884	921	LAC	940	1165	LAC
12	1962	621	766	LAC	771	788	LAC	752	1193	LAC
13	1963	838	1054	1085	1207	1146	LAC	1055	1668	LAC
14	1964	658	853	758	892	705	772	LAC	1320	1014
15	1965	592	723	810	760	817	878	LAC	1062	1016
16	1966	1082	1166	830	1190	1009	924	894	1276	961
17	1967	358	426	545	542	555	525	564	748	714
18	1968	ABS	1163	992	ABS	ABS	ABS	975	2265	ABS
19	1969	768	893	826	883	786	711	ABS	1347	1093
20	1970	482	673	868	890	638	766	LAC	1232	1095
21	1971	851	956	944	1109	1059	931	LAC	1468	1061

Fig. 56 : Sélection des données pour le comblement des lacunes

Dans votre tableau original, vous allez sélectionner la matrice de données à compléter.

En première ligne, on doit trouver les noms des variables et en première colonne, les noms des observations. La figure 56 illustre une telle sélection de huit variables et douze observations. Cette plage étant sélectionnée, vous allez la copier et la coller dans la feuille "Modbt.xls" qui se trouve dans le répertoire HYDROLAB. Une fois ce collage fait, il vous suffit de compléter les quatre premières lignes du tableau comme le montrent les figures 57 et 58.

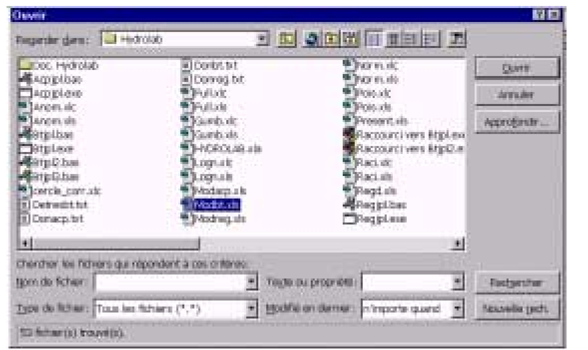
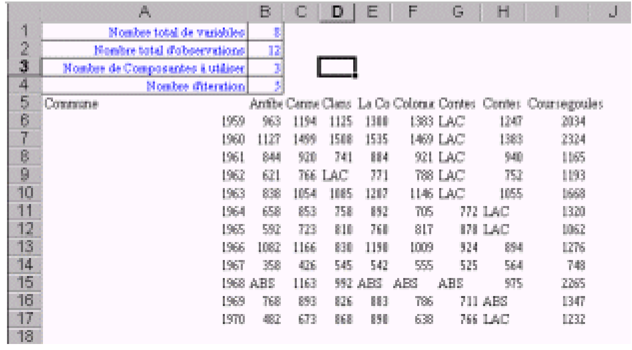


Fig. 57 : Ouverture de la feuille modèle "Modbt.xls"



	A	B	C	D	E	F	G	H	I
1	Nombre total de variables	8							
2	Nombre total d'observations	12							
3	Nombre de Composantes à utiliser	3							
4	Nombre d'iteration	5							
5	Commune	Antibe	Canne	Clans	La Co	Coloma	Contes	Contes	Coursegoules
6	1959	963	1194	1125	1300	1383	LAC	1247	2034
7	1960	1127	1499	1508	1535	1469	LAC	1383	2324
8	1961	844	920	741	884	921	LAC	940	1165
9	1962	621	766	LAC	771	788	LAC	752	1193
10	1963	838	1054	1085	1207	1146	LAC	1055	1668
11	1964	658	853	758	892	705	772	LAC	1320
12	1965	592	723	810	760	817	878	LAC	1062
13	1966	1082	1166	830	1190	1009	924	894	1276
14	1967	358	426	545	542	555	525	564	748
15	1968	ABS	1163	992	ABS	ABS	ABS	975	2265
16	1969	768	893	826	883	786	711	ABS	1347
17	1970	482	673	868	890	638	766	LAC	1232
18									

Fig. 58 : La feuille modèle prête à être utilisée

2.7.2 *Exécution de Bouche trous*

Lorsque la feuille modèle est complétée telle qu'à la figure 58, il suffit de lancer la commande "Bouche trous" du menu déroulant HYDROLAB.

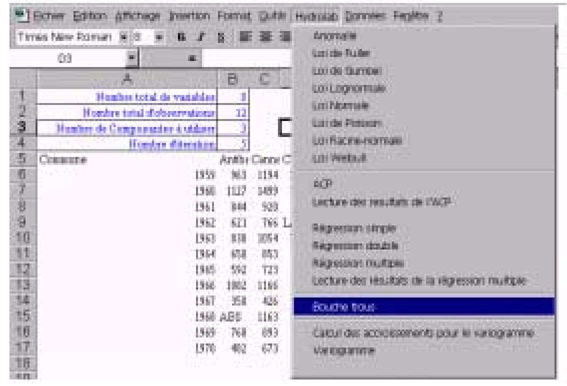
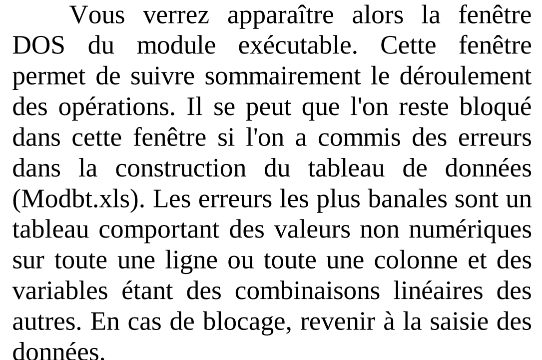


Fig. 59 : Lancement de "bouche trous"

Une fois lancée la commande, deux boites de dialogue vous seront soumises :

A la première boite de dialogue il convient de répondre oui. (la macro utilise un fichier temporaire déjà existant "Donbt.txt".

Pour la seconde boite de dialogue il convient de répondre non, en effet, HYDROLAB est conçu pour tourner indifféremment dans les versions 4, 5 et 97 d'Excel, les fichiers sont donc tous à sauvegarder en version 4 (compatibilité ascendante)



Vous verrez apparaître alors la fenêtre DOS du module exécutable. Cette fenêtre permet de suivre sommairement le déroulement des opérations. Il se peut que l'on reste bloqué dans cette fenêtre si l'on a commis des erreurs dans la construction du tableau de données (Modbt.xls). Les erreurs les plus banales sont un tableau comportant des valeurs non numériques sur toute une ligne ou toute une colonne et des variables étant des combinaisons linéaires des autres. En cas de blocage, revenir à la saisie des données.

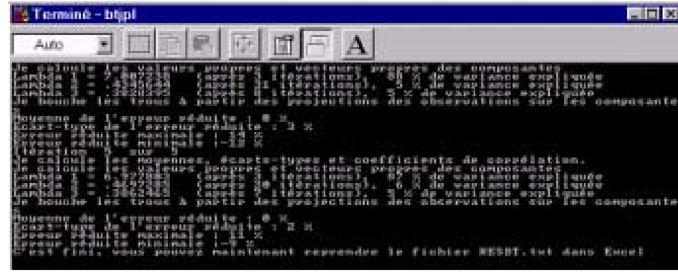


Fig. 60 : Fenêtre du module exécutable de "Bouche trous"

Suivant les versions, cette fenêtre se ferme automatiquement ou sinon il vous suffit de cliquer sur la croix en haut à droite de la fenêtre (figure 60)

2.7.3 *Visualisation des résultats de Bouche trous*

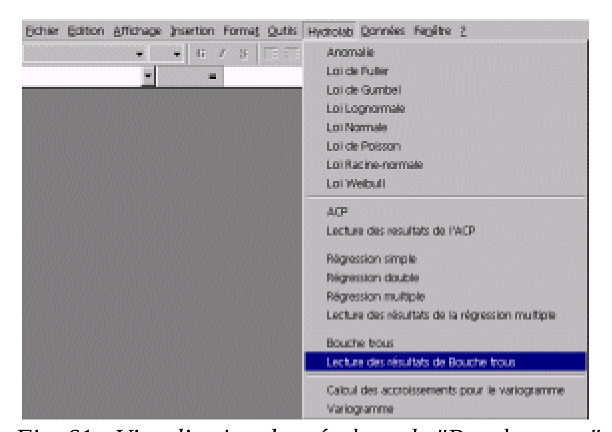


Fig. 61 : Visualisation des résultats de "Bouche trous"

Une fois la commande "Bouche trous" exécutée, vous pourrez visualiser les résultats en lançant la commande "Lecture des résultats de Bouche trous" du menu déroulant d'HYDROLAB.

Pendant quelques instant la commande va récupérer les résultats du module exécutable et les remettre sous la forme de deux tableaux Excel nommés "DetResbt.xls" et "Resbt.xls". Ces fichiers seront écrasés à chaque nouvelle utilisation. Une boite de dialogue vous demandera l'autorisation de remplacer l'ancienne version, il convient de répondre par oui.

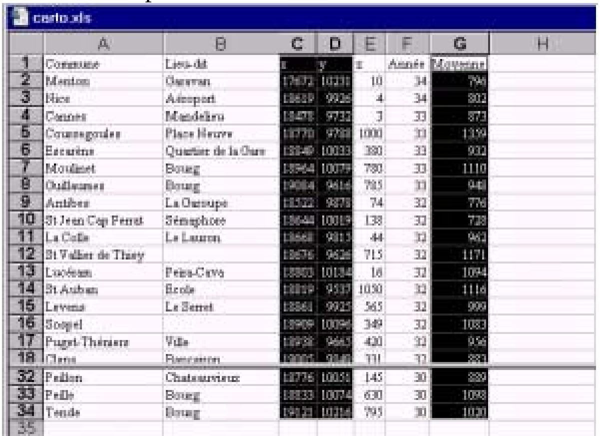
Les résultats sont représentés d'une part dans un fichier contenant la matrice originale (Resbt.xls) mais où les lacunes sont remplacées par leurs estimations et d'autre part dans un fichier donnant le détail des évolutions des estimations d'une itération à l'autre (Detresbt.xls). Ce fichier ne sert qu'à vérifier la convergence du processus.

2.8 Variographie

La mise au point de cette commande s'explique par l'utilisation fréquente du logiciel SURFER dans les synthèses cartographiques. Ce logiciel est interfacé avec Excel, mais il ne comprend pas de module d'analyse géostatistique. L'objectif de cette commande est donc d'effectuer une modélisation des variogrammes expérimentaux de façon à guider le choix dans les paramètres de krigeage de SURFER. Ce texte s'inspire principalement des travaux de C. OBLED 1986 .

2.8.1 Sélection des données

Chaque point de mesure est caractérisé par ses coordonnées cartésiennes et par la valeur mesurée en ce point de la fonction dont on veut étudier la fonction structure. Il convient donc de sélectionner dans une feuille Excel les coordonnées ainsi que les mesures sous la forme de trois colonnes dont la première contient les abscisses, la deuxième les ordonnées et la troisième la variable. Dans la première ligne on sélectionnera les noms de ces colonnes. La figure 62 illustre la sélection des coordonnées des postes pluviométriques des Alpes maritimes disposant d'au moins trente années de mesures, et de la précipitation moyenne interannuelle.



carto.xls

	A	B	C	D	E	F	G	H
1	Commune	Lieu-dit	x	y	z	Année	Moyenne	
2	Menton	Garavan	17672	10231	10	34	796	
3	Nice	Aéroport	18619	9926	4	34	802	
4	Cannes	Mandelieu	18478	9732	3	33	873	
5	Coursegoules	Place Neuve	18770	9788	1000	33	1359	
6	Escarène	Quartier de la Gare	18849	10033	380	33	932	
7	Moulinet	Bourg	18964	10079	780	33	1110	
8	Guillaumes	Bourg	19084	9616	785	33	948	
9	Antibes	La Garoupe	18522	9878	74	32	776	
10	St Jean Cap Ferrat	Sémaphore	18644	10019	138	32	728	
11	La Colle	Le Lauron	18668	9815	44	32	962	
12	St Vallier de Thiey		18676	9626	715	32	1171	
13	Lucéram	Peira-Cava	18803	10184	16	32	1094	
14	St Auban	Ecole	18819	9537	1050	32	1116	
15	Levens	Le Serret	18861	9925	565	32	999	
16	Sospel		18909	10096	349	32	1083	
17	Puget-Théniers	Ville	18938	9665	420	32	956	
18	Clans	Bancairon	[illegible]	[illegible]	331	32	[illegible]	
32	Peillon	Chateauvieux	18776	10051	145	30	889	
33	Peille	Bourg	18833	10074	630	30	1098	
34	Tende	Bourg	19121	10216	795	30	1020	
35								

Fig. 62 : sélection des données pour la construction du variogramme

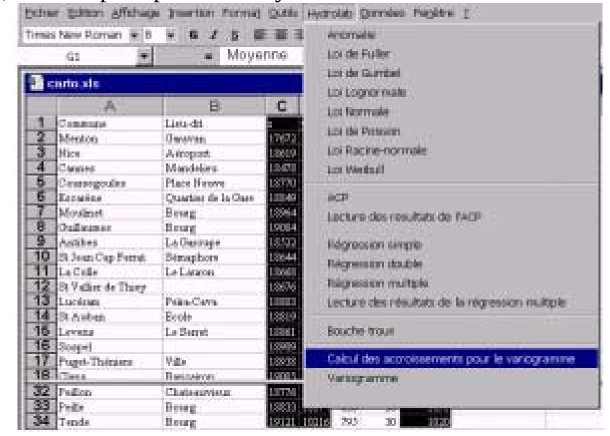


Fig. 63 : Lancement de la commande " Calcul des accroissements pour le variogramme

2.8.2 Constitution des couples

Cette sélection effectuée, on lance la commande "calcul des accroissements pour le variogramme" comme le montre la figure 63. Cette commande permet de constituer la feuille "Vario.xls" dans laquelle sont récapitulés tous les couples de points de mesure envisageables avec les distances entre points, les azimuts relatifs et les accroissements des mesures.

	A	B	C	D	E	F	W	X	Y	Z	AA	AB
1	N°	x	y	Moyenne	Nb de valeurs	33	Cat.	effec.	h	g(h)	Az0	0
2	1	17672	10231	796	Moyenne	979,22					Tolérance	90
3	2	18619	9926	802	Variance	17684					Pas	87,53
4	3	18478	9732	873	Ecart-type	132,98					Distance max. d'analyse	1751
5	4	18770	9788	1359	Nb de couples	528						
6	5	18849	10033	932	Distance max. absolue	1750,5						
7	6	18964	10079	1110	Distance max. dans la direction	1750,5						
8	7	19084	9616	948								
9	8	18522	9878	776								
10	9	18644	10019	728								
11	10	18668	9815	962								
12	11	18676	9626	1171								
13	12	18803	10184	1094								
14	13	18819	9537	1116								
15	14	18861	9925	999								
16	15	18909	10096	1083								
17	16	18938	9665	956								
18	17	19005	9849	883								
19	18	19080	9870	1045								
20	19	19085	9937	1144			1	19	64,023	5017,3		
21	20	19098	9816	976								
22	21	19106	9682	1047								
23	22	19272	9659	978								
24	23	19308	9603	1079								
25	24	18689	9854	1006								
26	25	18868	9709	976								
27	26	19083	10205	1033								
28	27	19222	9755	953								
29	28	18715	9974	772								
30	29	18718	10014	828								
31	30	18750	9931	884								
32	31	18776	10051	889								
33	32	18833	10074	1098								
34	33	19121	10216	1020								
35												

Fig. 64 : Description de la feuille "Vario.xls"

Comme le montre la figure 64, en colonne E et F sont indiqués les principales caractéristiques de l'échantillon à savoir sa taille, la moyenne, la variance et l'écart-type des valeurs, le nombre de couples possible et la distance maximale entre deux points de mesure. Dans les colonnes AA et AB, vous avez la possibilité de jouer sur les paramètres déterminants les couples à prendre en compte dans la constitution du variogramme expérimental.

La première précaution est de prendre maximale d'analyse largement inférieure à la distance maximale (la moitié, voir le tiers). Ayant choisi cette distance, vous pouvez également fixer le pas de distances séparant les couples en classes. Enfin, il est possible de ne prendre en compte que les couples tels que leur azimut relatif soit compris entre :

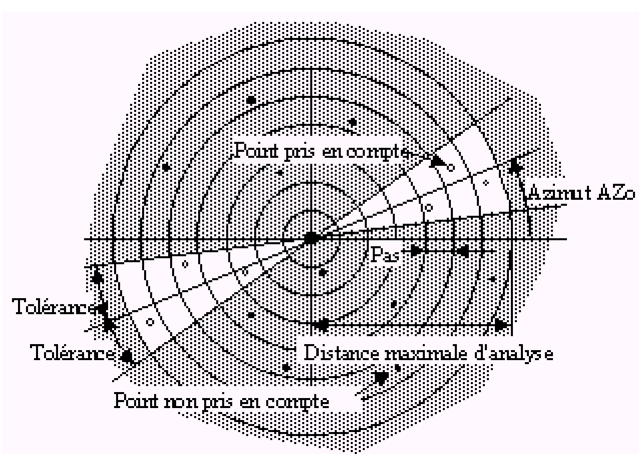


Fig. 65 : Paramètres d'un variogramme anisotrope

AZ_0 - tolérance et AZ_0 + tolérance

Ainsi que l'indique la figure 65, il est donc possible de construire des variogrammes expérimentaux anisotropes. Attention pour l'azimut, la convention est de donner l'angle en degré, à partir de l'axe des abscisses et en tournant dans le sens trigonométrique (sur la figure on a sensiblement Azo=30° et une tolérance de 15°).

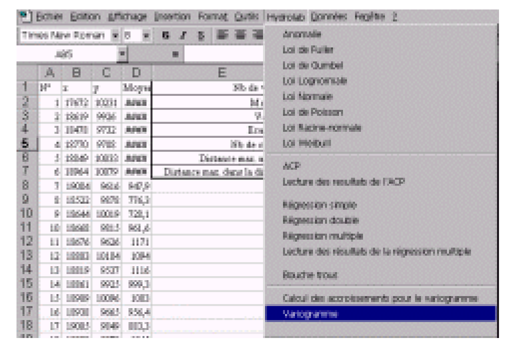
A priori, les paramètres proposés sont ceux d'un variogramme omnidirectionnel (tolérance de 90°) calculé sur la moitié du champ. Le pas d'exploration des distances est le 20ème de la moitié du champ (20 classes de distances).

2.8.3 Calcul du variogramme expérimental

E	F	W	X	Y	Z	AA	AB
Nb de valeurs	33	Cat.	effec.	h	g(h)	Az0	10
Moyenne	979,218826					Tolérance	45
Variance	17683,7755					Pas	40
Ecart-type	132,980358					Distance max. d'analyse	400
Nb de couples	528						
Distance max. absolue	1730,52564						
Distance max. dans la direction	399,025063	1	1	39,5600809	83,3632042		

Fig. 66 : Choix des caractéristiques du variogramme.

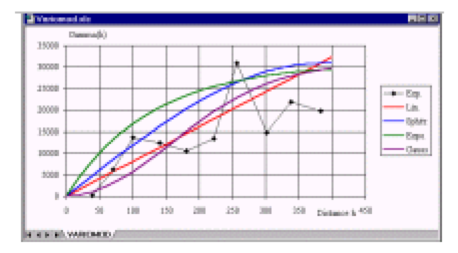
Une fois choisie les paramètres du variogramme (ici Fig. 66, on a pris une tolérance de plus ou moins 45° autour d'une direction moyenne à 10°, la distance maximale d'analyse est de 400 et l'on a des classes de 40), on peut lancer la construction des variogrammes expérimentaux et théoriques comme l'indique la fig. 67.


Fig. 67 :Lancement du calcul des variogrammes théoriques et expérimental.

Variomod.xls

	A	B	C	D	E	F	Q	R	S	T	U
1	Az0	10	Cat.	effec.	h moy.	Exp.	Type de modèle	Portée	Palier	Pépite	Erreur type
2	Tolérance	45	1	1	39,56	83,3632	Linéaire	382,998425	30888,19573	0	6524,540945
3	Pas	40	2	7	70,598	6058,854	Sphérique	382,998425	30888,19573	0	8123,266387
4	Distance max. d'analyse	400	3	11	99,988	13752,15	Exponentiel	382,998425	30888,19573	0	9094,465602
5			4	17	140,36	12418,37	Gaussien	382,998425	30888,19573	0	6768,724436
6			5	16	188,66	18636,27					
7			6	25	221,81	13415,81					
8			7	28	256,84	30888,2					
9			8	14	301,39	14855,02					
10			9	25	330,26	23011,62					
11			10	23	382,99	19742,69					
12											

Fig. 68 Feuille récapitulative des caractéristiques des variogrammes


Fig. 69 Représentation graphique des variogrammes théoriques et expérimental

Le lancement de cette commande permet la construction de deux feuilles. La première comme l'illustre la figure 68 récapitule les différents éléments constitutifs du variogramme expérimental (pour chaque catégorie on trouve successivement l'effectif, la distance moyenne et le γ(h) expérimental), et dans la partie droite, les caractéristiques des variogrammes théoriques (type, portée palier et pépite), la colonne "erreur type" caractérise les écarts entre le variogramme expérimental et les théoriques. Cette colonne est un guide pour l'optimisation des paramètres du variogramme théorique à ajuster. La seconde feuille Fig. 69, présente les résultats précédents sous forme graphique.

2.8.4 Ajustement du variogramme théorique

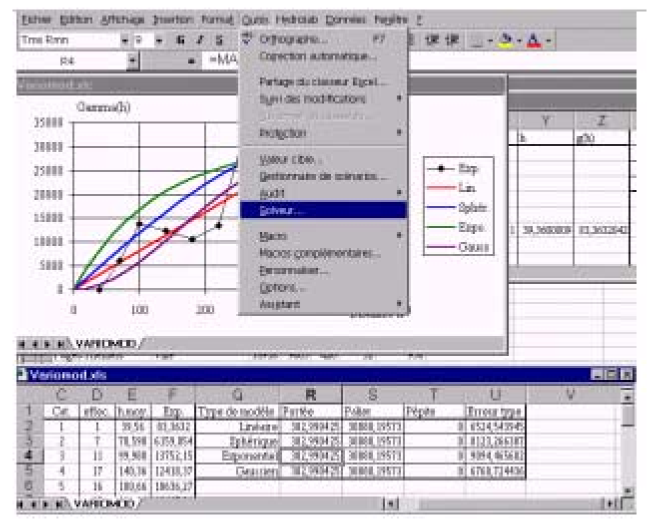


Fig. 70 Lancement du Solveur d'Excel.

Avant d'ajuster les paramètres il convient tout d'abord de choisir un des quatre modèles théoriques proposés. Dans l'exemple de la fig. 69, il paraît raisonnable de retenir un modèle gaussien (l'effectif de la première catégorie n'est que de 1 mais la suite semble indiquer un comportement très lisse à l'origine). Il nous reste alors que trois paramètres à ajuster. La première méthode consiste à modifier manuellement les paramètres et à tâtonner au regard des graphiques.

La seconde méthode consiste à utiliser le Solveur d'Excel comme l'illustre la Fig. 70.

On obtient alors une boite de dialogue dans laquelle on va préciser ce que l'on veut optimiser, les paramètres susceptibles d'être modifiés et d'éventuelles contraintes.

La fig. 71 montre l'ajustement à un modèle gaussien obtenu par minimisation de l'erreur type correspondante (cellule U5) en faisant varier les portée, palier et pépite (cellules R5 à T5) sous la seule contrainte que la pépite soit positive (cellule T5).
L'optimisation est effectuée en cliquant sur le bouton "résoudre".

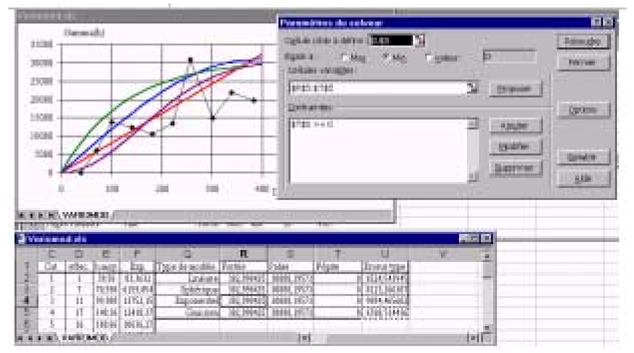


Fig. 71 Choix des paramètres du Solveur.

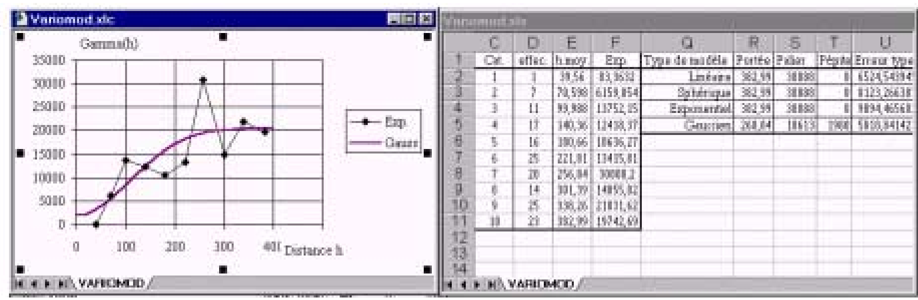


Fig. 72 Résultat de l'ajustement par le Solveur.

Dans la mesure où l'on avait choisi le modèle gaussien, nous avons allégé le graphique de la Fig. 72 en supprimant les autres modèles. Il suffit pour cela de cliquer sur la courbe correspondante et de taper la touche "supprimer".

2.8.5 Construction d'un nouveau variogramme

Si sur les même données, on désire construire un nouveau variogramme expérimental (changement de direction, de tolérance, de pas de distance…) il convient de fermer les deux feuilles "variomod.xls" et "variomod.xlc". On revient alors à la feuille "vario.xls" dans laquelle on indiquera les nouveaux paramètres de construction du variogramme expérimental puis on procédera à nouveau comme indiqué au paragraphe 2.7.3.

2.8.6 Définitions et rappels théoriques

Pour plus de renseignements on peut se référer à l'article de C. OBLED (1986).

Les quatre modèles proposés correspondent aux définitions suivantes :

Modèle linéaire : $\gamma(h) = \text{Pépite} + h\dfrac{\text{Palier}}{\text{Portée}}$

Modèle Sphérique : $\begin{cases} \gamma(h) = \text{Pépite} + \text{Palier}\left[\dfrac{3}{2}\dfrac{h}{\text{Portée}} - \dfrac{1}{2}\dfrac{h^3}{\text{Portée}^3}\right] & \text{Si } h \langle \text{ Portée} \\ \gamma(h) = \text{Pépite} + \text{Palier} & \text{Si } h \rangle \text{ Portée} \end{cases}$

Modèle Exponentiel : $\gamma(h) = \text{Pépite} + \text{Palier}\left(1 - e^{-\frac{3h}{\text{Portée}}}\right)$

Modèle gaussien : $\gamma(h) = \text{Pépite} + \text{Palier}\left(1 - e^{-\left(\frac{1.732\ h}{\text{Portée}}\right)^2}\right)$

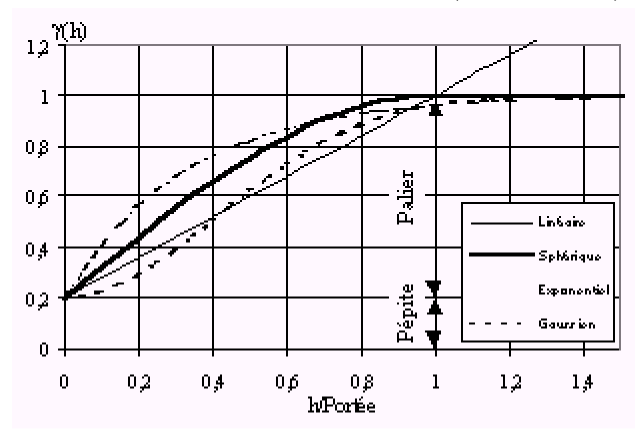


Fig. 73 : Variogrammes théoriques utilisés

Comme le montre la Fig. 73, les définitions des variogrammes théoriques sont telles que lorsque la distance h est égale à la portée, le terme γ(h) est sensiblement égal à Pépite + Palier.

Ce que nous appelons "portée" est donc ce que d'autres appelle parfois "portée pratique" ou encore "distance de décorrélation".

Ce choix paraîtra peut être curieux à des puristes mais il est justifié par son côté pratique. En effet, la "portée" que vous ajusterez sous HYDROLAB, correspond au paramètre "Range" nécessaire au krigeage sous SURFER.

3 Les fonctions personnalisées

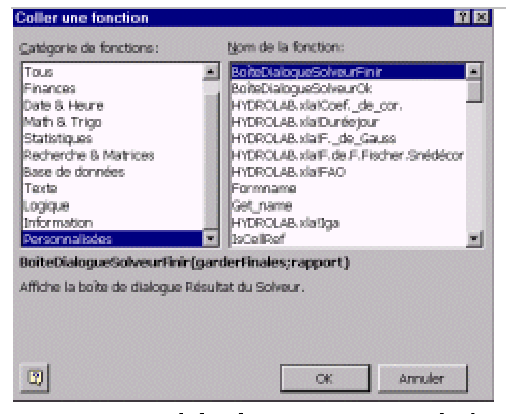


Fig. 74 : Appel des fonctions personnalisées

Ces fonctions sont accessibles par le menu "Insertion" où en choisissant "Fonction" on accède à la boite de dialogue de la Fig. 74. En choisissant dans les catégories de fonctions celles nommées "Personnalisées", on a accès en colonne de gauche à différentes fonctions d'HYDROLAB. Elles sont rangées par ordre alphabétique et précédées du préfixe "HYDROLAB.xla!".

3.1 Fonctions statistiques

3.1.1 U._de_Gauss

Cette fonction permet d'évaluer la valeur de la variable réduite de Gauss U correspondant à une fréquence au non dépassement F. Elle ne possède qu'un seul argument d'entrée F qui doit être compris entre 0 et 1.
Pour les justifications on peut se reporter au paragraphe 2.1.2.

	A	B
1	F	U
2	0,01	=HYDROLAB.xla!U._de_Gauss(A2)
3	0,02	=HYDROLAB.xla!U._de_Gauss(A3)
4	0,05	=HYDROLAB.xla!U._de_Gauss(A4)
5	0,1	=HYDROLAB.xla!U._de_Gauss(A5)
6	0,2	=HYDROLAB.xla!U._de_Gauss(A6)
7	0,5	=HYDROLAB.xla!U._de_Gauss(A7)
8	0,8	=HYDROLAB.xla!U._de_Gauss(A8)
9	0,9	=HYDROLAB.xla!U._de_Gauss(A9)
10	0,95	=HYDROLAB.xla!U._de_Gauss(A10)
11	0,98	=HYDROLAB.xla!U._de_Gauss(A11)
12	0,99	=HYDROLAB.xla!U._de_Gauss(A12)
13		

	A	B
1	F	U
2	0,01	-2,326785333
3	0,02	-2,054188589
4	0,05	-1,64521144
5	0,10	-1,281728757
6	0,20	-0,841456717
7	0,50	1,01007E-07
8	0,80	0,841456717
9	0,90	1,281728757
10	0,95	1,64521144
11	0,98	2,054188589
12	0,99	2,326785333
13		

Fig. 75 : Exemple d'utilisation de la fonction U_de_Gauss
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

3.1.2 F._de_Gauss

Cette fonction est l'inverse de la précédente où connaissant une valeur de la variable réduite U, on peut obtenir une estimation de sa fréquence au non dépassement F.
Pour les justifications on peut se reporter au paragraphe 2.1.2.

	A	B
1	U	F
2	-2,32678533255897	=HYDROLAB.xla!F._de_Gauss(A2)
3	-2,0541885887219	=HYDROLAB.xla!F._de_Gauss(A3)
4	-1,64521144014382	=HYDROLAB.xla!F._de_Gauss(A4)
5	-1,28172875650271	=HYDROLAB.xla!F._de_Gauss(A5)
6	-0,841456717354784	=HYDROLAB.xla!F._de_Gauss(A6)
7	1,01006675468085E-07	=HYDROLAB.xla!F._de_Gauss(A7)
8	0,841456717354784	=HYDROLAB.xla!F._de_Gauss(A8)
9	1,28172875650271	=HYDROLAB.xla!F._de_Gauss(A9)
10	1,64521144014381	=HYDROLAB.xla!F._de_Gauss(A10)
11	2,0541885887219	=HYDROLAB.xla!F._de_Gauss(A11)
12	2,32678533255897	=HYDROLAB.xla!F._de_Gauss(A12)
13		

	A	B
1	U	F
2	-2,326785333	0,010011689
3	-2,054188589	0,019819783
4	-1,64521144	0,049778124
5	-1,281728757	0,100097926
6	-0,841456717	0,20014752
7	1,01007E-07	0,50000004
8	0,841456717	0,79985248
9	1,281728757	0,899902074
10	1,64521144	0,950221876
11	2,054188589	0,980180217
12	2,326785333	0,989988311
13		

Fig. 76 : Exemple d'utilisation de la fonction F_de_Gauss
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

3.1.3 F. de F. de Fisher Snedecor

Cette fonction donne la fréquence au non dépassement d'une valeur F de Fisher Snedecor pour des degrés de liberté ν1 et ν2.

La variable $x = \dfrac{F^{1/3}\left(1-\dfrac{2}{9\nu_2}\right)-\left(1-\dfrac{2}{9\nu_1}\right)}{\sqrt{\dfrac{2}{9\nu_1}+F^{2/3}\dfrac{2}{9\nu_2}}}$ suit une loi de Gauss et c'est ainsi que l'on attribue une fréquence au F de Fisher Snedecor. Pour plus de détail on se référera à ABRAMOWITZ (1964), page 947.

	A	B
1	Valeur de F ayant	
2	n2/n1	1
3	5	6,61
4	10	4,96
5	15	4,54
6	20	4,35
7	25	4,24
8	30	4,17
9	Probabilité calculé	
10	n2/n1	1
11	5	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B3;B$2;$A3)
12	10	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B4;B$2;$A4)
13	15	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B5;B$2;$A5)
14	20	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B6;B$2;$A6)
15	25	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B7;B$2;$A7)
16	30	=HYDROLAB.xla!F. de F.Fischer.Snédécor(B8;B$2;$A8)

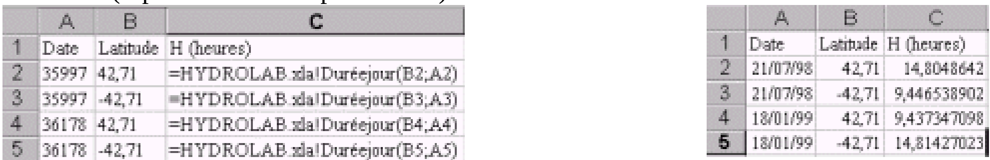
	A	B	C	D	E	F
1	Valeur de F ayant la probabilité 0,05 d'être dépassée					
2	n2/n1	1	3	5	8	12
3	5	6,61	5,41	5,05	4,82	4,68
4	10	4,96	3,71	3,33	3,07	2,91
5	15	4,54	3,29	2,9	2,64	2,48
6	20	4,35	3,1	2,71	2,45	2,28
7	25	4,24	2,99	2,6	2,34	2,16
8	30	4,17	2,92	2,53	2,27	2,09
9	Probabilité calculée					
10	n2/n1	1	3	5	8	12
11	5	0,951	0,949	0,949	0,949	0,949
12	10	0,952	0,950	0,950	0,950	0,950
13	15	0,952	0,951	0,950	0,950	0,950
14	20	0,952	0,951	0,950	0,950	0,950
15	25	0,953	0,951	0,950	0,951	0,950
16	30	0,953	0,951	0,950	0,951	0,950

Fig. 77 : Exemple d'utilisation de la fonction F. de F. de Fisher Snedecor
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

3.2 Fonctions liées à l'ETP et l'ETR

3.2.1 Durée astronomique du jour (Duréejour)

Cette fonction donne la durée du jour exprimée en heures en fonction de la date et de la latitude exprimée en degré et dixièmes (+ pour le nord et - pour le sud).



	A	B	C
1	Date	Latitude	H (heures)
2	35997	42,71	=HYDROLAB.xla!Duréejour(B2;A2)
3	35997	-42,71	=HYDROLAB.xla!Duréejour(B3;A3)
4	36178	42,71	=HYDROLAB.xla!Duréejour(B4;A4)
5	36178	-42,71	=HYDROLAB.xla!Duréejour(B5;A5)

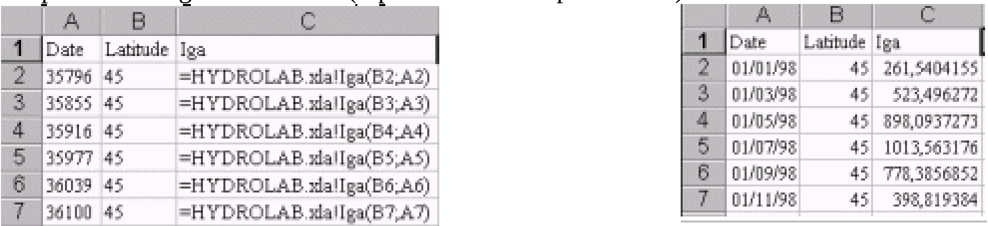
	A	B	C
1	Date	Latitude	H (heures)
2	21/07/98	42,71	14,8048642
3	21/07/98	-42,71	9,446538902
4	18/01/99	42,71	9,437347098
5	18/01/99	-42,71	14,81427023

Fig. 78 : Exemple d'utilisation de la fonction Duréejour
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe II présentant un extrait de LABORDE (1997).

3.2.2 Radiation globale théorique (Iga)

Cette fonction donne la radiation globale théorique (en calories par cm^2 et par jour) en fonction de la date et de la latitude exprimée en degré et dixièmes (+ pour le nord et - pour le sud).



	A	B	C
1	Date	Latitude	Iga
2	35796	45	=HYDROLAB.xla!Iga(B2;A2)
3	35855	45	=HYDROLAB.xla!Iga(B3;A3)
4	35916	45	=HYDROLAB.xla!Iga(B4;A4)
5	35977	45	=HYDROLAB.xla!Iga(B5;A5)
6	36039	45	=HYDROLAB.xla!Iga(B6;A6)
7	36100	45	=HYDROLAB.xla!Iga(B7;A7)

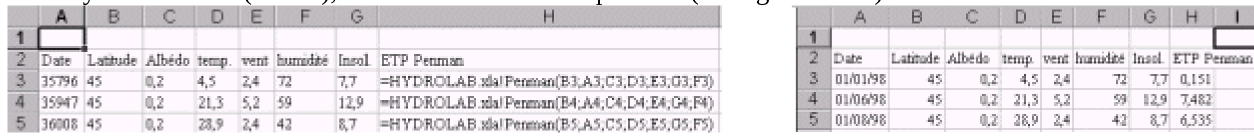
	A	B	C
1	Date	Latitude	Iga
2	01/01/98	45	261,5404155
3	01/03/98	45	523,496272
4	01/05/98	45	898,0937273
5	01/07/98	45	1013,563176
6	01/09/98	45	778,3856852
7	01/11/98	45	398,819384

Fig. 79 : Exemple d'utilisation de la fonction Duréejour
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe II présentant un extrait de LABORDE (1997).

3.2.3 ETP Penman originale (Penman)

Cette fonction permet d'évaluer l'ETP Penman exprimée en mm/jour en fonction de la date, de la latitude (en degré et dixièmes), de l'humidité relative (en %), de l'insolation journalière (en heures et dixièmes), du vent moyen à 2 mètres (en m/s), de l'Albédo et de la température (en degré Celsius).



	A	B	C	D	E	F	G	H
1								
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol.	ETP Penman
3	35796	45	0,2	4,5	2,4	72	7,7	=HYDROLAB.xla!Penman(B3;A3;C3;D3;E3;G3;F3)
4	35947	45	0,2	21,3	5,2	59	12,9	=HYDROLAB.xla!Penman(B4;A4;C4;D4;E4;G4;F4)
5	36008	45	0,2	28,9	2,4	42	8,7	=HYDROLAB.xla!Penman(B5;A5;C5;D5;E5;G5;F5)

	A	B	C	D	E	F	G	H	I
1									
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol.	ETP Penman	
3	01/01/98	45	0,2	4,5	2,4	72	7,7	0,151	
4	01/06/98	45	0,2	21,3	5,2	59	12,9	7,482	
5	01/08/98	45	0,2	28,9	2,4	42	8,7	6,535	

Fig. 80 : Exemple d'utilisation de la fonction Penman
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe II présentant un extrait de LABORDE (1997).

3.2.4 ETP Penman modifiée FAO (FAO)

Cette fonction permet d'évaluer l'ETP Penman modifiée par la FAO exprimée en mm/jour en fonction de la date, de la latitude (en degré et dixièmes), de l'humidité relative (en %), de l'insolation journalière (en heures et dixièmes), du vent moyen à 2 mètres (en m/s), de l'Albédo et de la température (en degré Celsius).

	A	B	C	D	E	F	G	H
1	Altitude =	755						
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol.	FAO
3	35796	45	0,2	4,5	2,4	72	7,7	=HYDROLAB.xla!FAO(B3;A3;C3;D3;E3;G3;F3)
4	35947	45	0,2	21,3	5,2	59	12,9	=HYDROLAB.xla!FAO(B4;A4;C4;D4;E4;G4;F4)
5	36008	45	0,2	28,9	2,4	42	8,7	=HYDROLAB.xla!FAO(B5;A5;C5;D5;E5;G5;F5)

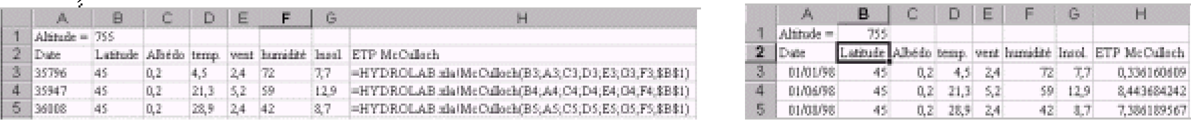
	A	B	C	D	E	F	G	H
1	Altitude =	755						
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol.	FAO
3	01/01/98	45	0,2	4,5	2,4	72	7,7	0,271
4	01/06/98	45	0,2	21,3	5,2	59	12,9	8,144
5	01/08/98	45	0,2	28,9	2,4	42	8,7	7,265

Fig. 81 : Exemple d'utilisation de la fonction FAO
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe II présentant un extrait de LABORDE (1997).

3.2.5 ETP Penman modifiée Mc Culloch (McCulloch)

Cette fonction permet d'évaluer l'ETP Penman modifiée par Mc Culloch exprimée en mm/jour en fonction de la date, de la latitude (en degré et dixièmes), de l'humidité relative (en %), de l'insolation journalière (en heures et dixièmes), du vent moyen à 2 mètres (en m/s), de l'Albédo, de l'altitude (en m) et de la température (en degré Celsius).



	A	B	C	D	E	F	G	H
1	Altitude =	755						
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol	ETP McCulloch
3	35796	45	0,2	4,5	2,4	72	7,7	=HYDROLAB.xla!McCulloch(B3;A3;C3;D3;E3;G3;F3;B1)
4	35947	45	0,2	21,3	5,2	59	12,9	=HYDROLAB.xla!McCulloch(B4;A4;C4;D4;E4;G4;F4;B1)
5	36008	45	0,2	28,9	2,4	42	8,7	=HYDROLAB.xla!McCulloch(B5;A5;C5;D5;E5;G5;F5;B1)

	A	B	C	D	E	F	G	H
1	Altitude =	755						
2	Date	Latitude	Albédo	temp.	vent	humidité	Insol.	ETP McCulloch
3	01/01/98	45	0,2	4,5	2,4	72	7,7	0,336160609
4	01/06/98	45	0,2	21,3	5,2	59	12,9	8,443684242
5	01/08/98	45	0,2	28,9	2,4	42	8,7	7,386189567

Fig. 82 : Exemple d'utilisation de la fonction McCulloch
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe II présentant un extrait de LABORDE (1997).

3.2.6 Evapotranspiration réelle (ETR)

Cette fonction permet d'évaluer l'évapotranspiration réelle sur un pas de temps quelconque, en connaissant la pluie et l'ETP sur ce même pas de temps, la RFU maximale possible et l'état en début de pas de temps de cette RFU. On suppose qu'à tout instant le rapport de l'évapotranspiration réelle instantanée à l'évapotranspiration potentielle instantanée est égal au rapport de l'état instantané de la RFU à sa capacité maximale.

	A	B	C	D	E
1	RFUmax	**100**			RFUinit
2	Mois	ETP	Pluie	ETR	**15**
3	1	25	45	=HYDROLAB.xla!ETR(B1;E2;B3;C3)	=MIN(E2+C3-D3;B1)
4	2	12	51	=HYDROLAB.xla!ETR(B1;E3;B4;C4)	=MIN(E3+C4-D4;B1)
5	3	39	75	=HYDROLAB.xla!ETR(B1;E4;B5;C5)	=MIN(E4+C5-D5;B1)
6	4	45	56	=HYDROLAB.xla!ETR(B1;E5;B6;C6)	=MIN(E5+C6-D6;B1)
7	5	86	25	=HYDROLAB.xla!ETR(B1;E6;B7;C7)	=MIN(E6+C7-D7;B1)
8	6	123	5	=HYDROLAB.xla!ETR(B1;E7;B8;C8)	=MIN(E7+C8-D8;B1)
9	7	156	0	=HYDROLAB.xla!ETR(B1;E8;B9;C9)	=MIN(E8+C9-D9;B1)
10	8	215	0	=HYDROLAB.xla!ETR(B1;E9;B10;C10)	=MIN(E9+C10-D10;B1)
11	9	165	0	=HYDROLAB.xla!ETR(B1;E10;B11;C11)	=MIN(E10+C11-D11;B1)
12	10	57	15	=HYDROLAB.xla!ETR(B1;E11;B12;C12)	=MIN(E11+C12-D12;B1)
13	11	14	79	=HYDROLAB.xla!ETR(B1;E12;B13;C13)	=MIN(E12+C13-D13;B1)
14	12	18	152	=HYDROLAB.xla!ETR(B1;E13;B14;C14)	=MIN(E13+C14-D14;B1)

	A	B	C	D	E
1	RFUmax	**100**			RFUinit
2	Mois	ETP	Pluie	ETR	**15**
3	1	25	45	8,5	51,5
4	2	12	51	8,8	93,7
5	3	39	75	38,8	100,0
6	4	45	56	45,0	100,0
7	5	86	25	65,9	59,1
8	6	123	5	43,9	20,1
9	7	156	0	15,9	4,2
10	8	215	0	3,7	0,5
11	9	165	0	0,4	0,1
12	10	57	15	3,6	11,5
13	11	14	79	6,8	83,7
14	12	18	152	17,8	100,0

Fig. 83 : Exemple d'utilisation de la fonction ETR
(à gauche affichage des fonctions, à droite affichage des valeurs numériques)

Pour plus de détail on se reportera à l'annexe III présentant un extrait de LABORDE (1998).

3.3 Fonctions des coordonnées géographiques

HYDROLAB permet de passer des coordonnées géographiques (longitude et latitude) aux coordonnées des projections Lambert fréquemment utilisées. Longitudes et latitudes sont exprimées en ° et dixièmes (les longitudes est sont comptées positivement), les coordonnées Lambert sont exprimées en mètres.

Six systèmes de projection Lambert sont disponibles : France (zone 1 Nord, zone 2 Centre, zone 3 Sud, zone 4 Corse) et Algérie (zone Nord Algérie et Sud Algérie).

3.3.1 Passage des coordonnées géographiques aux Lambert

Il suffit comme pour les autres fonctions de les faire apparaître dans le menu "Insertion / Fonction / Fonctions personnalisées". Les algorithmes nous ont été aimablement fournis par J.P. PIRAT du service de géodésie et de nivellement de **l'Institut Géographique National**, ils sont ceux en vigueur au 15/03/1996.

	A	B	C	D	E
1	Lieux	Longitude	Latitude	XLambert (m)	Ylambert (m)
2	Environs de Lille	2,99	50,65	646172,3115	328108,5675
3	Environs de Lyon	4,8	45,7	791791,5966	80741,52637
4	Environs de Marseille	5,2	43,5	831479,0653	137367,6587
5	Environs d'Ajaccio	8,8	41,9	535769,3018	390860,5389
6	Environs d'Alger	3	37	526698,523	410969,6613
7	Environs de Ouargla	5,5	31,9	764823,6258	148346,2411

	A	B	C	D	E
1	Lieux	Longitude	Latitude	XLambert (m)	Ylambert (m)
2	Environs de Lille	2,99	50,65	=HYDROLAB.xla!XL1($C2;$B2)	=HYDROLAB.xla!YL1($C2;$B2)
3	Environs de Lyon	4,8	45,7	=HYDROLAB.xla!XL2($C3;$B3)	=HYDROLAB.xla!YL2($C3;$B3)
4	Environs de Marseille	5,2	43,5	=HYDROLAB.xla!XL3($C4;$B4)	=HYDROLAB.xla!YL3($C4;$B4)
5	Environs d'Ajaccio	8,8	41,9	=HYDROLAB.xla!XL4($C5;$B5)	=HYDROLAB.xla!YL4($C5;$B5)
6	Environs d'Alger	3	37	=HYDROLAB.xla!XLNA($C6;$B6)	=HYDROLAB.xla!YLNA($C6;$B6)
7	Environs de Ouargla	5,5	31,9	=HYDROLAB.xla!XLSA($C7;$B7)	=HYDROLAB.xla!YLSA($C7;$B7)

Fig. 84 : Exemples d'utilisation des fonctions géographiques
(en bas affichage des fonctions, en haut affichage des valeurs numériques)

3.3.2 Passage des coordonnées Lambert aux géographiques

Le passage inverse des coordonnées Lambert aux coordonnées géographiques, peut être réalisé aisément en utilisant le Solveur d'Excel. La figure 85 illustre la démarche qui consiste à minimiser la distance entre deux points, le premier connus par ses coordonnées Lambert, le second dépendant de la longitude et de la latitude.

	A	B	C	D	E
1		Point connu	Coordonnées Géo.		Lambert estimé
2	XLambert (m)	646172	Longitude	1	=HYDROLAB.xla!XL1(D3;D2)
3	Ylambert (m)	328108	Latitude	1	=HYDROLAB.xla!YL1(D3;D2)
4			Distance	=((B2-E2)^2+(B3-E3)^2)^0,5	
5					

Fig. 85 : Préparation de la feuille

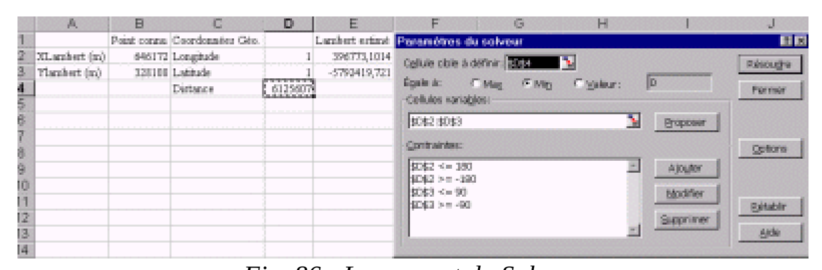


Fig. 86 : Lancement du Solveur

	A	B	C	D	E
1		Point connu	Coordonnées Géo.		Lambert estimé
2	XLambert (m)	646172	Longitude	2,988964	646099,3267
3	Ylambert (m)	328108	Latitude	50,64972	328076,2946
4			Distance	79,28839	
5					
6					

Fig. 87 : Résultats obtenus

REFERENCES BIBLIOGRAPHIQUES UTILISEES

ABRAMOWITZ M. & STEGUN I. - "Handbook of mathematical functions" - United states department of commerce - 1964 - Washington - 1046 pages.

BRUNET MORET Y. - "*Etude de quelques lois statistiques utilisées en hydrologie*" - Cahiers ORSTOM série hydrologie - 1969 - Vol. 6 - N° 3 - PP. 3-100.

C.T.G.R.E.F. - "*Utilisation de quelques tests en hydrologie et calcul de l'intervalle de confiance*" - Informations techniques du C.T.G.R.E.F. - 1978 - Cahier 31 - N° 2 - 4 pages.

COLIN E.- "*Etude des moments et autres caractéristiques de quelques lois hydrologiques*" - Note interne du C.T.G.R.E.F. - 1977 - Antony - 49 pages.

COLIN E., MICHEL C. et OBERLIN G. - "*Etude statistique d'évenements hydrologiques observés sur un nombre réduit d'années*" - Note interne du C.T.G.R.E.F. - 1978 - Antony - 16 pages.

DURAND R. - "Estimation du rayonnement global à partir de la durée d'insolation" - Annales Agronomiques de l'INRA - 1974 - Paris -

LABORDE J. P. - "Eléments d'hydrologie de surface" - Cours polycopié de l'Université de Nice Sophia Antipolis - 1998 - Nice - 195 pages.

LABORDE J. P. - "Notice explicative de la carte des évapotranspirations potentielles de l'Algérie du Nord" - Projet de Coopération algéro-allemande n°94 21 83 5 - 1997 - GTZ-ANRH Alger - 41 pages.

LEBART L., MORINEAU A. et FENELON J. P. - "*Traitement des données statistiques*" - Dunod - 1979 - Paris - 510 pages.

LEGRAS J. - "*Méthodes et techniques de l'analyse numérique*" - Dunod - 1971 - Paris - 350 pages.

MIQUEL J. - "*Guide pratique d'estimation des probabilités de crues*" - Eyrolles - 1984 - Paris - 160 pages.

OBLED C. - "*Introduction au krigeage*" - Actes des deuxièmes journées hydrologiques de l'ORSTOM - 1986 - Montpellier -

ROCHE M. - "*hydrologie de surface*" - Gauthier Villars - 1963 - Paris - 430 pages.

YEVIJEVICH V. - "*Probability and statistics in hydrology*" - Water Resources Publications - 1982 - Littelton, Colorado - 302 pages.

ANNEXES

Annexe I : Comblement des lacunes dans des séries pluviométriques

Bien des études de synthèse sur les hauteurs de pluie annuelles, nécessitent de connaître les précipitations sur différentes stations mais sur une même période. C'est le cas par exemple pour l'établissement de cartes pluviométriques moyennes interannuelles, où il est vivement recommandé de travailler sur une période commune trentenaire. Malheureusement les nv postes pluviométriques n'ont pas tous été observés durant les no années de la période de référence et la matrice des données se présente donc généralement ainsi :

	x(1,1)	**inconnu**	...	x(1,j)	...	x(1,nv)
	x(2,1)	x(2,2)	...	x(2,j)	...	**inconnu**
	...	...	...	...	...	...
[x] =	x(i,1)	x(i,2)	...	x(i,j)	...	x(i,nv)
	inconnu	...	...	**inconnu**	...	...
	...	**inconnu**	...	...	...	...
	x(no,1)	x(no,2)	...	x(no,j)	...	x(no,nv)

Si le nombre de poste nv est grand, il est quasiment impossible de "boucher" les trous un par un par des techniques de régression. Le choix des variables explicatives est vite inextricable et de plus les erreurs accidentelles ou systématiques qui ne manquent pas d'affecter certaines observations, sont ainsi répétées. Nous proposons donc une méthode qui permets de combler rapidement et simplement les lacunes avec la partie la plus fiable de ce tableau de données.

Nous calculerons tout d'abord les moyennes expérimentales Mxo(j) et les écart-types expérimentaux Sxo(j) des précipitations pour chaque station j et sur les seules années réellement observées :

[Mxo] =	Mxo(1)	Mxo(2)	...	Mxo(j)	...	Mxo(nv)
[Sxo] =	Sxo(1)	Sxo(2)	...	Sxo(j)	...	Sxo(nv)

On peut alors évaluer la matrice des valeurs centrées réduites : $u(i,j) = \frac{x(i,j) - Mxo(j)}{Sxo(j)}$

et calculer pour chaque observation la moyenne des valeurs centrées réduites disponibles : $\bar{u}(i)$

	u(1,1)	**inconnu**	...	u(1,j)	...	u(1,nv)	$\bar{u}(1)$
	u(2,1)	u(2,2)	...	u(2,j)	...	**inconnu**	$\bar{u}(2)$
	...	...	...	...	...	...	…
[u] =	u(i,1)	u(i,2)	...	u(i,j)	...	u(i,nv)	$\bar{u}(i)$
	inconnu	...	...	**inconnu**	...	...	…
	...	**inconnu**	...	...	...	...	…
	u(no,1)	u(no,2)	...	u(no,j)	...	u(no,nv)	$\bar{u}(no)$

On peut alors dans une première étape remplacer chaque valeur inconnue x(i,j) par une valeur $xe_0(i,j) = \frac{\bar{u}(i) - Mxo(j)}{\sigma xo(j)}$ correspondant à la variable réduite moyenne pour cette observation. On obtient alors une première matrice complète [xo] : $\mathbf{xe_0(i,j)}$

	x(1,1)	$\mathbf{xe_0(1,2)}$	...	x(1,j)	...	x(1,nv)
	x(2,1)	x(2,2)	...	x(2,j)	...	$\mathbf{xe_0(2,nv)}$
	...	...	...	...	...	...
[xo] =	x(i,1)	x(i,2)	...	x(i,j)	...	x(i,nv)
	$\mathbf{xe_0(i+1,1)}$	...	...	$\mathbf{xe_0(i+1,j)}$	...	...
	...	$\mathbf{xe_0(no-1,2)}$	...	...	...	...
	x(no,1)	x(no,2)	...	x(no,j)	...	x(no,nv)

Nous effectuons alors une A.C.P. sur cette matrice [xo] et obtenons les projections des variables $[a_1]$ et des observations $[c_1]$ sur les k seules premières composantes principales pouvant avoir une signification physique :

$$[a_1] = \begin{vmatrix} a_1(1,1) & a_1(1,2) & \ldots & a_1(1,j) & \ldots & a_1(1,nv) \\ a_1(2,1) & a_1(2,2) & \ldots & a_1(2,j) & \ldots & a_1(2,nv) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_1(k,1) & a_1(k,2) & \ldots & a_1(k,j) & \ldots & a_1(k,nv) \end{vmatrix} \qquad [c_1] = \begin{vmatrix} c_1(1,1) & c_1(1,2) & \ldots & c_1(1,k) \\ c_1(2,1) & c_1(2,2) & \ldots & c_1(2,k) \\ \ldots & \ldots & \ldots & \ldots \\ c_1(i,1) & c_1(i,2) & \ldots & c_1(i,k) \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ c_1(no,1) & c_1(no,2) & \ldots & c_1(no0,k) \end{vmatrix}$$

Ainsi que les matrices des moyennes et écarts-types :

$$\begin{matrix} [Mx_1] = \\ [Sx_1] = \end{matrix} \begin{vmatrix} Mx_1(1) & Mx_1(2) & \ldots & Mx_1(j) & \ldots & Mx_1(nv) \\ Sx_1(1) & Sx_1(2) & \ldots & Sx_1(j) & \ldots & Sx_1(nv) \end{vmatrix}$$

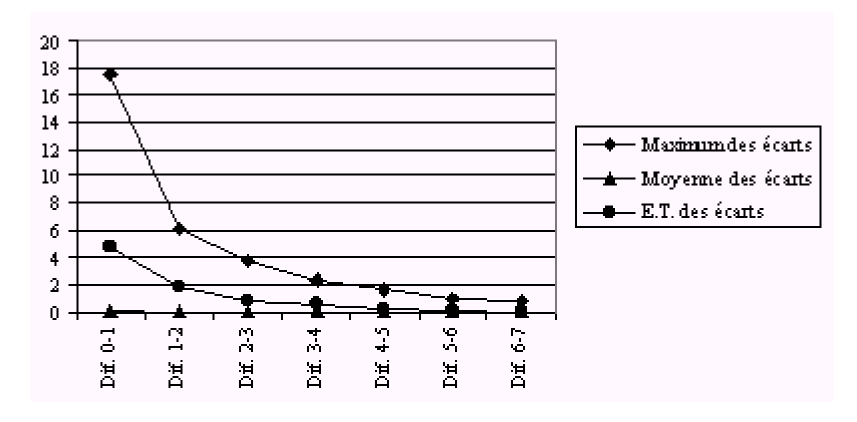
Il est alors possible de reconstituer chaque observation manquante x(i,j) par une nouvelle valeur estimée $xe_1(i,j)$:

$$xe_1(i,j) = Mx_1(j) + Sx_1(j) * \{ a_1(1,j)\, c_1(i,1) + a_1(2,j)\, c_1(i,2) + \ldots + a_1(k,j)\, c_1(i,k) \}$$

Cette estimation n'est pas très correcte puisque l'A.C.P. a été effectuée sur une matrice "bouchée" à partir de moyennes interannuelles, cependant $xe_1(i,j)$ est une meilleure estimation que $xe_0(i,j)$ puisqu'elle tient compte des observations aux autres stations pour cette année j. On peut donc réitérer le processus en remplaçant dans la matrice [xo] chaque $xe_0(i,j)$ par les $xe_1(i,j)$ adaptés. On obtient ainsi une nouvelle matrice $[x_1]$:

$$[x_1] = \begin{vmatrix} x(1,1) & \mathbf{xe_1(1,2)} & \ldots & x(1,j) & \ldots & x(1,nv) \\ x(2,1) & x(2,2) & \ldots & x(2,j) & \ldots & \mathbf{xe_1(2,nv)} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ x(i,1) & x(i,2) & \ldots & x(i,j) & \ldots & x(i,nv) \\ \mathbf{xe_1(..,1)} & \ldots & \ldots & \mathbf{xe_1(i,j)} & \ldots & \ldots \\ \ldots & \mathbf{xe_1(..,2)} & \ldots & \ldots & \ldots & \ldots \\ x(no,1) & x(no,2) & \ldots & x(no,j) & \ldots & x(no,nv) \end{vmatrix}$$

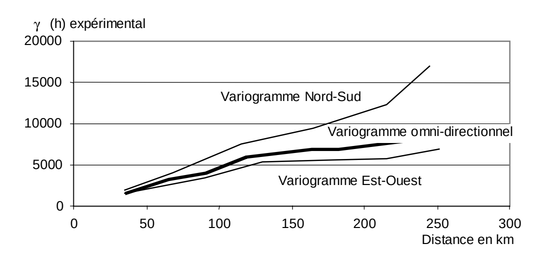
On recommence alors une A.C.P. sur la matrice $[x_1]$ permettant d'obtenir de nouvelles matrices $[c_2]$, $[a_2]$, $[Mx_2]$ et $[Sx_2]$ d'où l'on tirera de nouvelles estimations $xe_2(i,j)$.



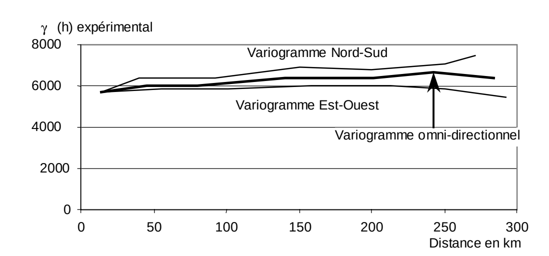
A chaque itération nous modifions les estimations pour les observations manquantes. A titre d'exemple, on citera l'étude de 120 postes pluviométriques d'Algérie sur une période commune de 60 ans où nous avions 1854 stations-années manquantes sur 7200 (25% de "trous"). Nous avons à chaque fois noté la moyenne et l'écart-type des modifications, ainsi que la plus forte modification en valeur absolue. Ces résultats sont reportés sur le graphe ci-contre.

On y constate que toutes les caractéristiques des écarts d'une itération à l'autre suivent sensiblement une exponentielle décroissante. Ici comme dans d'autres tentatives on constate que le processus est stabilisé dès la 6ème ou 7éme itérations.

La difficulté réside dans le choix du nombre de composantes principales à utiliser dans les reconstitutions. Nous préconisons bien sûr, de n'utiliser que les composantes ayant une signification physique. La pluie ayant à priori une structure spatiale, on peut préjuger que seules les composantes ayant également une structure spatiale sont physiquement significatives. Ainsi le choix des composantes à utiliser se basera sur l'analyse du variogramme des projections des variables sur les composantes :



Comportement à l'origine de la troisième composante



Comportement à l'origine de la quatrième composante

Dans cet exemple algérien, les trois premières composantes ont une structure spatiale nette (mais anisotrope), par contre la quatrième composante est pépitique pure. On ne prendra donc en compte dans les reconstitutions que les trois premières composantes.

Annexe II : Estimation de l'évapotranspiration potentielle par la méthode de Penman et ses dérivées

II.1 Estimation de l'E.T.P. par la méthode de Penman originale

Penman propose d'évaluer l'ETP à partir d'un bilan énergétique simple :
Rn = A + S + E * L

Rn représente la radiation nette reçue au sol ;
A représente le flux de chaleur au bénéfice de l'atmosphère ;
S le flux de chaleur résultant des échanges thermiques avec le sol ;
E le flux évaporé
et L la chaleur latente.

La formule développée prend la forme suivante :

$$ETP = \left\{ \frac{Iga}{59}(1-a)\left(0{,}18 + 0{,}62\,\frac{h}{H}\right) - \frac{\sigma}{59}T^4\left(0{,}56 - 0{,}08\sqrt{e}\right)\left(0{,}10 + 0{,}90\,\frac{h}{H}\right)\right\} \frac{\frac{F'_T}{\gamma}}{1 + \frac{F'_T}{\gamma}} + \frac{0{,}26}{1 + \frac{F'_T}{\gamma}}\left(e_w - e\right)\left(1 + 0{,}54\,V\right)$$

ETP	évapotranspiration potentielle en mm/j ;
Iga	radiation solaire directe en l'absence d'atmosphère exprimée en cal/cm^2/j ;
a	albédo de la surface évaporante, prise ici pour de la végétation à 0,2 ;
h	durée réelle d'insolation en heures et dixièmes du jour considéré ;
H	durée maximale possible d'insolation en heures et dixièmes pour ce jour ;
σ	constante de STEFAN-BOLTZMAN soit 1,18 10^{-7} cal/cm^2/jour/°K
T	température moyenne journalière de l'air sous abri exprimée en degrés Kelvin (T = t + 273 si l'on mesure la température t en ° Celsius) ;
e	tension moyenne journalière de la vapeur d'eau mesurée sous abri et exprimée en millibars ;
F'_T	pente de la courbe de tension de vapeur saturante pour la température de l'air T ;
γ	constante psychrométrique ;
e_w	tension maximale possible de la vapeur d'eau, exprimée en millibars, pour la température T ;
V	Vitesse moyenne journalière du vent mesuré à 2 mètres au-dessus de la surface évaporante et exprimée en m/s.

Ces différents paramètres intervenant dans l'évaluation de l'ETP proviennent :

- de mesures directes sur le terrain pour h, T, et V ;

- de mesures indirectes pour e, puisque e est estimée à partir de l'humidité relative Hr en %, et de la température t en ° Celsius par la relation :

$$e = 0{,}061\ H_r\ 10^{\left(7{,}5\frac{t}{238+t}\right)} \quad \text{(en millibars)}$$

- de constantes physiques bien connues σ, F'_T, γ et e_w :

$$\sigma = 1{,}18\ 10^{-7}\ \text{cal/cm}^2\text{/jour/°K}$$

$$F'_T = \frac{6,149}{T} \exp\left(19,511 \frac{t}{T}\right) \left(\frac{6463}{T} - 3,927\right)$$

(F'_T en millibars par ° Celsius, t en ° Celsius, T en ° Kelvin)

$$\gamma = 0,6605 - 0,826\ 10^{-4}\ z$$

(z altitude du lieu en mètres, γ en millibar par ° Celsius,
en fait, on prendra $\gamma \approx$ Cte =0,66)

$$e_w = 6,107\ 10^{\left(7,5 \frac{t}{238+t}\right)}$$ (e_W en millibars, t en ° Celsius)

- de constantes physiques dépendant de la latitude L du lieu et de la déclinaison D du soleil en fonction de la date J . Les formules que nous donnons ici sont issues de R. DURAND ("Estimation du rayonnement global à partir de la durée d'insolation", Annales Agronomiques de l'INRA, 1974). Nous avons comparé les résultats de ces formules avec les tables fournies par la F.A.O. (Bulletin N° 24) et y avons constaté des écarts généralement inférieurs à 1% ou du même ordre que le nombre de décimales données (5% au maximum).
  la déclinaison est donnée par :
  D = 0,0066241 + 0,406149 sin [0,0172029 (J - 81,95)]
  + 0,006675 sin [0,0344057 (J - 42,85)]
  + 0,003009 sin [0,0516086 (J - 21,42)]
  + 0,000149 sin [0,0688115 (J - 17,57)]
  (D s'exprime en radians et J est le numéro du jour dans l'année,
  de 1 pour le 1er janvier à 365 (ou 366) pour le 31 Décembre)

$$H = 7,6394 \left\{ \text{Arc} \cos[-tg(L)\ tg(D)] + \frac{0,01065}{\sqrt{\cos^2(L) - \sin^2(D)}} \right\}$$

(H est en heures, et L représente la latitude en radians)

$$Iga = 914,54 * \left\{ \sin(L) \sin(D)\ \text{Arc} \cos[-tg(L)\ tg(D)] + \sqrt{\cos^2(L) - \sin^2(D)} \right\} * \{1 + 0,033 \cos(0,0172\ J)\}$$

(Iga s'exprime en calories par cm^2 et par jour)

- d'une constante physique à évaluer sommairement en fonction de l'état de surface (les mesures sur le terrain sont extrêmement rares) :
  nappes d'eau a ≈ 0,05 à 0,07
  cultures vertes a ≈0,2
  herbes et savane a ≈0,22
  forêt a ≈0,11
  roches a ≈0,16
  sables clairs a ≈0,26
  ...

II.2 Estimation par la méthode de Penman modifiée F.A.O.

La formule de Penman a été modifiée par de nombreux auteurs de façon à en simplifier l'écriture ou en adapter les paramètres à des régions particulières. Parmi ces formules modifiées on utilise fréquemment au Maghreb la formule dite de Penman–FAO que nous donnons ici avec les paramètres recommandés pour les régions sèches et arides.

$$ETP = \left\{ \frac{Iga}{59} (1-a) \left(0,25 + 0,45 \frac{h}{H}\right) - \frac{\sigma}{59} T^4 (0,56 - 0,08 \sqrt{e}) \left(0,10 + 0,90 \frac{h}{H}\right) \right\} \frac{\frac{F'_T}{\gamma}}{1 + \frac{F'_T}{\gamma}} + \frac{0,26}{1 + \frac{F'_T}{\gamma}} (e_w - e)(1 + 0,864\,V)$$

On y retrouve les mêmes variables que citées précédemment, mais seuls quelques paramètres changent :

$(0,25 + 0,45 \frac{h}{H})$ au lieu de $(0,18 + 0,62 \frac{h}{H})$

$(1 + 0,864\,V)$ au lieu de $(1 + 0,54\,V)$

II.3 Estimation par la méthode de Penman modifiée par Mc Culloch

Parmi les modifications apportées à cette formule, citons également celles de Mc Culloch. En effet cette modification a été considérée comme bénéfique sous le climat marocain. Cette formule se présente sous la forme suivante :

$$ETP = \left\{ \frac{Iga}{59} (1-a) \left(0,29 \cos(L) + 0,52 \frac{h}{H}\right) - \frac{\sigma}{59} T^4 (0,56 - 0,08 \sqrt{e}) \left(0,10 + 0,90 \frac{h}{H}\right) \right\} \frac{\frac{F'_T}{\gamma}}{1 + \frac{F'_T}{\gamma}} + \frac{0,26}{1 + \frac{F'_T}{\gamma}} (e_w - e)(1 + 0,864\,V)\left(1 + \frac{z}{20\,000}\right)$$

On y retrouve les mêmes variables que dans la formule de Penman, mais la latitude L et l'altitude z du site (en mètres) sont introduites comme paramètres correcteurs :

$(0,29 \cos(L) + 0,52 \frac{h}{H})$ au lieu de $(0,18 + 0,62 \frac{h}{H})$

$(1 + 0,864\,V)(1 + \frac{z}{20\,000})$ au lieu de $(1 + 0,54\,V)$

Annexe III : Passage de l'Etp à l'ETR par un bilan tenant compte du stress hydrique de la végétation

La procédure classique de Thornwaite suppose que l'évapotranspiration réelle est égale à l'évapotranspiration potentielle, jusqu'à ce que la R.F.U. soit vide. Cependant avant que la R.F.U. soit vide, la végétation subi un stress hydrique et diminue son évapotranspiration. Il en résulte que si la part de la R.F.U. encore en eau diminue, la végétation n'évaporera pas toute l'E.T.P, mais une part réduite.

Nous supposerons qu'entre deux instants t_1 et t_2 distants de Δt il est tombé une quantité de pluie P que l'on pourra supposée d'intensité constante : i

$$i = \frac{P}{t_2 - t_1}$$

De même pendant cet intervalle de temps, on supposera que l'évapotranspiration potentielle E_p se produit à intensité constante : e_p

$$e_p = \frac{E_p}{t_2 - t_1}$$

Pour passer de l'intensité de l'évapotranspiration potentielle e_p à l'intensité de l'évapotranspiration réelle e_r , on supposera que cette dernière est proportionnelle à e_p et au rapport de l'état de la R.F.U. r(t), à sa capacité maximale R . L'évapotranspiration réelle er(t) est alors variable dans le temps :

$$e_r(t) = e_p \frac{r(t)}{R}$$

Sur un petit intervalle de temps dt, le bilan en eau de la réserve s'écrit :

$$d\{r(t)\} = i\,dt - e(t)\,dt$$

Cette relation étant valable tant que $r(t) \leq R$

On obtient ainsi l'équation différentielle suivante :

$$d\{r(t)\} = \left(i - \frac{e_p}{R} r(t) \right) dt$$

En posant $\alpha = e_p/R$ on obtient :

$$\frac{d\{r(t)\}}{r(t) - i/\alpha} = -\alpha\,dt$$

cette équation s'intègre aisément entre les instants t_1 et t :

$$\left[\mathrm{Ln}\{ r(t) - i/\alpha \} \right]_{t_1}^{t} = \left[-\alpha t \right]_{t_1}^{t}$$

$$\mathrm{Ln} \left(\frac{r(t) - i/\alpha}{r(t_1) - i/\alpha} \right) = -\alpha (t - t_1)$$

$$\boxed{r(t) = i/\alpha + \{r(t_1) - i/\alpha\} \mathbf{e}^{-\alpha(t-t_1)}}$$

a) cas où $r(t_2)$ reste inférieur à R

$$r(t_2) = i/\alpha + \{r(t_1) - i/\alpha\} \mathbf{e}^{-\alpha(t_2-t_1)}$$

$$\boxed{r(t_2) = \frac{P}{E_p} R + \{r(t_1) - \frac{P}{E_p} R\} \mathbf{e}^{-\frac{E_p}{R}}}$$

Pour calculer l'évapotranspiration réelle Er entre les instants t_1 et t_2, il suffit d'intégrer l'équation de $e_r(t)$:

$$Er = \int_{t_1}^{t_2} e_r(t)\,dt$$

$$Er = \int_{t_1}^{t_2} \frac{e_p}{R} r(t)\,dt = \int_{t_1}^{t_2} \alpha\, r(t)\,dt$$

En remplaçant r(t) par son expression calculée plus haut on obtient :

$$Er = \int_{t_1}^{t_2} [i + \{\alpha r(t_1) - i\}\mathbf{e}^{-\alpha(t-t_1)}]\, dt$$

$$Er = [\, i\, t\,]_{t_1}^{t_2} - \{r(t_1) - i/\alpha\}\, [\, \mathbf{e}^{-\alpha(t-t_1)}\,]_{t_1}^{t_2}$$

$$Er = i(t_2 - t_1) - \{r(t_1) - i/\alpha\}\{\mathbf{e}^{-\alpha(t_2-t_1)} - 1\}$$

$$\boxed{Er = P - \{r(t_1) - \frac{P}{E_p} R\}\{\mathbf{e}^{-\frac{E_p}{R}} - 1\}}$$

b) cas où $r(t_2)$ deviendrait supérieur R

On calcule alors le temps t' auquel r(t) atteint R :

$$r(t') = i/\alpha + \{r(t_1) - i/\alpha\}\mathbf{e}^{-\alpha(t'-t_1)} = R$$

$$\mathbf{e}^{-\alpha(t'-t_1)} = \frac{R - i/\alpha}{r(t_1) - i/\alpha}$$

$$t' = t_1 + 1/\alpha\ \mathrm{Ln}\ \{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\}$$

On calcule alors aisément Er :

$$Er = \int_{t_1}^{t'} e(t)\, dt + \int_{t'}^{t_2} e_p\, dt$$

$$Er = i(t' - t_1) - \{r(t_1) - i/\alpha\}\{\mathbf{e}^{-\alpha(t'-t_1)} - 1\} + e_p(t_2 - t')$$

$$Er = i/\alpha\ \mathrm{Ln}\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\} - \{r(t_1) - i/\alpha\}\{\mathbf{e}^{-\mathrm{Ln}\,\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\}} - 1\} + e_p(t_2 - t_1 - 1/\alpha\ \mathrm{Ln}\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\})$$

$$Er = i/\alpha\ \mathrm{Ln}\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\} - \{r(t_1) - i/\alpha\}\{\frac{R - i/\alpha}{r(t_1) - i/\alpha} - 1\} + e_p(t_2 - t_1) - e_p/\alpha\ \mathrm{Ln}\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\}$$

$$Er = (i/\alpha - e_p/\alpha)\ \mathrm{Ln}\{\frac{r(t_1) - i/\alpha}{R - i/\alpha}\} + e_p(t_2 - t_1) - \{R - r(t_1)\}$$

$$\boxed{Er = (\frac{P\,R}{E_p} - R)\ \mathrm{Ln}\left\{\frac{r(t_1) - \frac{P\,R}{E_p}}{R - \frac{P\,R}{E_p}}\right\} + E_p - \{R - r(t_1)\}}$$

c) Mise en œuvre pratique :

Entre les instants t_1 et t_2 distant de Δt, il est tombé une hauteur de pluie P, pendant cet intervalle de temps Δt, l'évapotranspiration potentielle est de E_p et à l'instant t_1, la réserve contenait $r(t_1)$.

On calcule tout d'abord :

$$r(t_2) = \frac{P}{E_p} R + \left\{r(t_1) - \frac{P}{E_p} R\right\} e^{-\frac{E_p}{R}}$$

Si le résultat est inférieur ou égal à R on aura :

$$\boxed{Er = P - \{r(t_1) - \frac{P}{E_p} R\}\{\mathbf{e}^{-\frac{E_p}{R}} - 1\} \quad \text{et} \quad r(t_2) = \frac{P}{E_p} R + \{r(t_1) - \frac{P}{E_p} R\}\mathbf{e}^{-\frac{E_p}{R}}}$$

Si le résultat est supérieur à R on aura :

$$\boxed{Er = (\frac{P\,R}{E_p} - R)\ \mathrm{Ln}\left\{\frac{r(t_1) - \frac{P\,R}{E_p}}{R - \frac{P\,R}{E_p}}\right\} + E_p - \{R - r(t_1)\} \quad \text{et} \quad r(t_2) = R}$$